

More realism still to come to oil

Last week's reduction of the price of OPEC oil is good news, but the delight of oil consumers is tinged with nervousness. They should go further, agreeing that their currencies need to be valued realistically.

If a man should sell a gadget which costs 50 cents to manufacture and charge \$34 for it, one of two things will happen. Potential customers will either be so outraged that they refuse to buy, or they will find some other source of supply. Since the first dramatic increase of the price of crude oil almost exactly ten years ago, the consumers dependent on imported oil have followed both courses of action. Conservation measures have reduced oil consumption, while the fraction of oil consumers' crude supplied by members of the Organization of Petroleum Exporting Countries (OPEC) has steadily declined, principally because of the arrival on the scene of oil from Britain (the North Sea) and Mexico. OPEC nevertheless managed to smile bravely at the end of last week's meeting in London, when a spokesman for the group said that the agreement to hold production at 17.5 million barrels a day (about 850 million tonnes a year) and to reduce the price to \$29 a barrel had been reached so as to offer to the consumers the advantages of price stability they have lacked in the past ten years. This is the spirit in which a street salesman selling turkeys on Christmas Eve will from time to time announce that he has now fixed a rock-bottom price, thereby inciting all within earshot to wonder where he will go next.

OPEC's immediate difficulty is to know whether the new price will stick. Winning some kind of agreement on the level of OPEC production, in itself something of a triumph for the diplomats, should help. But past experience shows that OPEC members needing dollars are frequently prepared to cheat, producing more oil than they have agreed and if necessary shading the price a little. It will be some time, some weeks, before it becomes clear what is happening in this curious market. The international oil companies, the only substantial buyers of crude oil in bulk, have a plain interest in waiting as long as possible before signing on for long-term supplies at this or even a lower price. Before then, the British National Oil Company (now called Britoil), which also has a need to keep production in the North Sea alive, will have to decide whether to cut its own selling price again. But further price reductions are unlikely to be huge. For they would only stimulate economic activity and thus increase the demand for oil. A further 20 per cent off the new OPEC price is the most that can be hoped for. But that would be a great boon for the consumers.

So why are the oil consumers not over the Moon with delight? Part of the explanation is natural conservatism — people have become used to oil at \$34 a barrel. A more serious problem is that afflicting Mexico — an oil producer outside OPEC and dependent on its oil revenues for repayments on its overseas debts, incurred in the expectation of much larger oil revenues than those now in prospect. The oil consumers whose banks originally made these loans (now partly covered by borrowing from the International Monetary Fund) should be sympathetic, recognizing how much they stand to gain because Mexico never joined the would-be cartel. British gloom about the falling price of oil is harder to understand. Last week, the British Chancellor of the Exchequer was telling the House of Commons that he might have to revise his new budget, and raise more taxes, if the price of oil should fall further, reducing in the process the government's revenue from oil taxation. (For the time being, the public purse has not suffered, for the value of sterling has followed down the price of oil, leaving taxation which is calculated on sterling prices relatively unchanged.) Generally, while nobody denies that the dramatic increase of the price of oil ushered in the long recession,

there is a sense of disbelief that the process can be reversed to any substantial degree.

That view is mistaken. Even the British economy is not so rigid that it cannot respond favourably to reductions of its input price. The fact that the upheaval of the past few weeks has also brought about a long overdue devaluation of sterling against the dollar and the deutschemark is a measure of what may now be possible. (The industrialists who were complaining loudly just a year ago that sterling was then overpriced have nevertheless been strangely quiet in the past few days.) The long wrangle last weekend between the West European finance ministers about the realignment of European currencies is a sign of how irrational are governments' attachment to "strong" pounds, francs or other denominations. Far from being the equivalents of virtues, the values of nations' currencies determine the success of both external trade and internal investment in innovation. (If a currency is overvalued, it is more prudent to export capital than to invest it in novel technology.) It thus makes no sense that most European governments should at present be spending large sums of public money on the stimulation of innovation (see page 279) but at the same time should be attached to some other value for their currency than the market dictates. OPEC's problems, fortunately, have shown that in the end realism prevails. □

Who goes to Oxbridge?

Even Britain's oldest universities are changing, but in ways that seem at best ill-considered.

For most British universities, the haunting problem of the past few years has been to adjust to a smaller budget, but the universities of Oxford and Cambridge have been struggling with a more daunting task — how to defend themselves against the charge that, by being centres of excellence in several academic fields, they are also seats of privilege and are thus a baleful influence on British society as a whole. The charge implies a serious threat, that some future government might intervene directly in the academic management of these universities. (Two of the three political parties that will contest the next general election promise just that.) But the more immediate danger is that the two universities will so expediently trim their sails to the political pressures that they will irreparably damage themselves and by extension, British higher education as a whole.

Change is needed — indeed, is overdue. But what change? The outstanding problem for both establishments is that of student entry. The days have long since gone when Oxford and Cambridge were for practical purposes private universities for fee-paying students, with a small but significant scholarship programme for the recruitment of the less well off. Since the Depression half a century ago, increasing proportions of Oxbridge students have been supported by public funds, at first by means of competitive public scholarships, since the Second World War by means of fees and maintenance awards which local authorities are required by law to pay to students winning places at approved universities. (Even the University of Buckingham, founded only seven years ago on a strictly private basis, now qualifies.) Even so, students from private secondary schools (called "public" by the British) are disproportionately



conspicuous at both places.

The explanation is straightforward. Both universities select students by means of an entrance examination of their own, originally the device by means of which scholarships were awarded. In the past few decades, private secondary schools have tumbled to it that their leaving students will have a better chance of getting to Oxbridge if they stay on for a few months after their normal school course, and the artful middle classes, always quick to make the most of whatever social services are available, have readily concluded that the competition is well worthwhile. But the publicly supported secondary schools are less well-equipped to provide this service for those among their students who might wish to go to Oxbridge, while many students who might nurse such ambitions are deterred by the reputation of both places for elitism of several kinds. As a consequence, neither university can claim that it recruits just those students who would most benefit from its distinctive way of teaching students. It is understandable that reflective Oxbridge academics should have sensed that unless they changed their ways, a future government might do it for them.

The reforms on which Oxbridge has so far embarked are unfortunately merely cosmetic. It is now several years since Oxford abandoned the requirement that entering students should know some Latin, and then the requirement that they should at least know something of a modern language was dropped in deference to the poverty of language teaching in publicly supported secondary schools. Now the general paper in the Oxford entrance examination, in many ways the most challenging (and entertaining) instrument of its kind, is also to be dropped. Meanwhile, the university (to its credit) has for several years been seeking out potential students from maintained secondary schools, offering them places on the basis of their performance in a future school-leaving examination.

The defects of all these devices are that they throw together into the same pot students who differ markedly in their academic preparation. As yet, Oxford has made no concessions to this circumstance by softening the curriculum most of its students follow. At Cambridge, where students can if they wish (and try hard) follow less specialized courses of study, attempts are similarly being made to cast the net wider, while a majority of the colleges awarding scholarships (whose nominal value is now negligible) seems ready to give up the practice in the hope that people will then set less store by success. In the same spirit might the organizer of a prizefight decide not to award a purse for fear that the participants might fight too hard, and damage each other.

Oxbridge cannot hope to escape political pressure as easily. The urgent need is that its selection procedures should be seen to be equitable. At present they are not, principally because selection depends to some extent on the quality (and quantity) of preparation by secondary schools and because it is recognized that secondary schools vary enormously in their academic capability. Some schools are frankly bad: why should their unlucky students be further penalized by being for practical purposes denied an Oxbridge place? As it happens, in the British system even the good schools are also educationally bad — indeed, they are often those that set the pace in that systematic perversion of general education that requires that students should rehearse at school what they intend to study at university or elsewhere in higher education. The result is that even Oxbridge is denied liberally educated students, while its selection procedures further accentuate the lopsidedness of the school curriculum.

The only way out of the dilemma is to change the system. So long as the school curriculum is as specialized as it is at present, and until Oxbridge decides to select students by lot, it will remain vulnerable to the charge that its students are predominantly those whose parents could afford to send them to a good secondary school. But the Oxbridge influence is powerful enough — that is part of the complaint — for them to be able to pave the way towards constructive reform. And the way is clear. Potential students should be selected a year earlier, and on the basis of general and not specialized attainment, and should then embark

on a longer university course, not the three-year scamper through higher learning which occupies no more than 72 weeks of a student's time. Such a procedure would allow a new student a chance to decide what kind of specialist he or she will be. It would also be a valuable demonstration to the rest of the British educational system that it is possible (as most other countries show) to arrange that arts and humanities students know what science is about and that even technical people can string together sentences in some language other than their own, or even write a literate scientific paper. □

Mr Reagan's dog-days

Is the United States presidency just unlucky, or is it accident-prone for some deeper reason?

PRESIDENT Reagan has had a bad beginning to 1983. The new Congress, with its preponderance of Democrats in the House of Representatives, is not of his choosing but might have been less unwelcome if his first eighteen months had been more obviously successful. But what has gone wrong this year cannot be blamed on the elected representatives but on the President and his entourage. Troubles such as those at the Environmental Protection Agency (for this week's instalment, see page 279) and, more seriously, at the Arms Control Agency, are the consequences of mistakes that should not have happened. Mrs Anne Burford would not have been forced out disowned if the government had had a policy on the environment. The danger that the Senate will not agree to appoint the President's nominated successor to Mr Eugene Rostow would not have arisen if Mr Rostow had not been fired precipitately and unnecessarily. The consequence is that the Administration is having to cobble together on the backs of envelopes policies in both important fields which are likely, for the sake of political palatability, to make more concessions to Congress than the government will be able to live with comfortably. The fable of the dog crossing the stream with a bone in its mouth refers.

In such circumstances, it is natural that sceptics should ask what will next go wrong. This year's science budget is one obvious worry. Superficially, the research agencies have done well, the National Science Foundation perhaps the most conspicuous among them. Part of the explanation is that the Administration's delight at the recruitment of Dr Edward Knapp as the foundation's new director led it, last November, to change its tack on federal support for science education. Two years after deciding that federal support for science education must vanish, the Administration wrote in something in its 1984 budget for this forgotten cause — and then found Congress even more anxious to up the ante. But since neither partner in this complicated Dutch auction yet knows how it would spend \$1 million, let alone fifty or a thousand times as much (the White House and Congress respectively) the prospect is either that the money will not be spent or the cause on which it has been lavished will be discredited.

These are the errors of commission. More worrying are the fields in much has been promised but nothing has been done. It is, for example, just a year since the President's Science Advisor Dr George Keyworth announced the formation of a council of distinguished scientists to provide the government with advice on scientific questions. Dr Keyworth was careful to emphasize that his new committee was not a re-creation of that venerable (and liberal) talking-shop, the President's Scientific Advisory Committee, but a kind of task-force of hard-working people who would get things done. Its first assignment was to study the problem of the national laboratories — are there too many, do they cost too much and what, in any case, are they for? Although people were plainly hoping for answers within the year, a new budget has come and gone with hardly anything having been said about this gigantic problem. Here again, it seems, the administration has built a machine to serve its purposes but has forgotten to turn the handle. No wonder that people suspect it does not know what to do about science education. And that all those who depend on its success at making its decisions stick are nervous. □

US Environmental Protection Agency

Another hit-list and more embarrassment

Washington

THE resignation of Environmental Protection Agency (EPA) administrator Mrs Anne Burford has hardly put an end to the agency's troubles (see *Nature* 10 March, p.95). Last week, as half a dozen congressional panels continued their investigations of the agency, new evidence came to light of apparent political manipulation in staff appointments and agency actions.

First, the EPA official named in connection with the so-called "hit-list" of scientific advisers, Louis Cordia, was asked to resign amid charges that he had destroyed files being sought under a Freedom of Information Act request and his own admission that he had prepared a second "hit list" of agency employees.

This list, the first few pages of which were found in Cordia's files by congressional investigators, contains comments from Cordia's "close advisers" — apparently industry and conservative groups — on current and prospective EPA employees. Many are couched in ideological terms. Roy Albert of New York University, who serves on EPA's Carcinogen Assessment Group, is called "unacceptable to this Administration"; other comments include "philosophically attuned, professionally acceptable", "supportive of the new Administration if the new Administration is for nuclear power, business and conservative interests", and

tion's transition team for EPA, and subsequently served in EPA's Office of Federal Activities. A covering memo on the document, marked "From: Lou", asks Sullivan and Perry: "please do not share the recommendations + staff lists with anyone. I would feel the backlash".

Meanwhile, John Hernandez, now acting EPA administrator, is feeling the heat of the congressional investigation. Hernandez admitted last week that he had allowed Dow Chemical Company officials to read a draft report on dioxin contamination from its plant in Michigan and to suggest changes that deleted any reference to Dow's responsibility. While denying that he himself ordered the changes in the report, Hernandez admitted that he urged EPA staff members on several occasions to consider Dow's suggestions.

Congress is also continuing its investigation of other instances in which EPA of-

ficials may have favoured political considerations or private interests over the advice of their own scientists. One EPA staff member told congressional investigators of an "election-track" policy of speeding up announcements of selected hazardous waste-site clean-ups last year in time to help Republican candidates in the November congressional and gubernatorial races. And at least two former EPA officials, Rita Lavelle, who was assistant administrator for solid waste, and James Sanderson, a personal friend of Burford's who served as a high-level consultant to the agency, are being investigated for possible conflicts of interest between their EPA actions and their previous (in Lavelle's case) or current (in Sanderson's case) private employers.

The congressional investigations may yet turn up more damaging evidence. But they have already done considerable damage to the efforts of Reagan's appointees to substitute ideological expedience for what many observers say could have been a well-founded and scientifically-based policy of deregulation. The major results of these efforts have been political embarrassment, a paralysed EPA and the alienation of its employees and scientific advisers.

Stephen Budiansky

UK industrial research

Budget brings good tidings

THE Department of Industry did well out of the British Government's budget for 1983-84, made public recently in the House of Commons. The department, as foreshadowed in last month's public expenditure plans, will have substantially increased funds to spend on support for industrial research and development — an extra £185 million over three years. As yet, however, it has not modified its spending plans to accommodate the urgent criticisms made by the House of Lords Select Committee on Science and Technology.

Most of the new money will be spent as part of the department's scheme called "Support for Innovation" — a grant-making programme which can meet part of the cost of development projects mounted by industrial companies, and which has prompted an embarrassingly large number of applications since it was devised in May last year by the amalgamation of earlier industrial support schemes.

One new feature of the programme is that funds will now be available to help companies to bridge the gap between the successful completion of a development project and the point at which a new product or process can be exploited. A total of £40 million has been earmarked for such uses.

Under the terms of the Support for Innovation scheme, grants of up to 33 1/3 per cent of the costs of approved research and development projects are offered to industrial companies. One of the most successful components of the scheme was

the Small Engineering Firms Investment Scheme, which helped small companies with grants for new capital equipment. That was closed ahead of schedule last year when the £30 million allocated to it had been fully committed: in 8 1/2 weeks, 1,757 companies had applied for assistance. A new version of the scheme, which will probably be similar to the old one, is to be relaunched shortly with £100 million of new money.

Other schemes to be expanded provide support for development of telecommunications equipment, software products and computer-aided design and manufacture. In addition, increased support of £20 million will be allocated to various advisory services. A loan guarantee scheme will be extended with an increased lending ceiling.

There is as yet no indication that support for individual projects will be extended above the 33 1/3 per cent level, or that the department will relax the rule that companies applying for support for a project must not have started work before applying, cited by the House of Lords committee as a cause of delay and thus of frustration. Moreover, it remains to be seen whether the department will heed the committee's complaint that to require applicant companies to prove that they would not undertake with their own resources the development projects for which they seek support is to ask them a Catch-22 question.

In other respects, the new support is likely to be welcomed by industry.

Tim Beardsley



"understood to have brought in many votes on 4 November", a reference to the date of the election.

The extent to which Cordia's comments were acted on is not clear. Congressional investigators are seeking the full 55-page document to find out. Cordia sent the document to Robert Perry, EPA's general counsel, and to William Sullivan, the former enforcement counsel, in July 1981, when he was on the staff of the conservative Heritage Foundation. Cordia had earlier served on the Reagan Administra-

Biotechnology

On the heels of interleukin-2

A MAJOR step nearer the commercial production of human interleukin-2 is announced today (see p.305) by Dr T. Taniguchi and his colleagues at the Tokyo Cancer Research Institute and the Ajinomoto Company. Production of this material has become the goal of many biotechnology companies, rather as was interferon two years ago.

Formerly known as T-cell growth factor, interleukin-2 can be produced naturally in the body in tiny amounts. When its activity was discovered a few years ago, it was claimed to have an important role in the growth in culture of T cells, the basis of cell-mediated immunity. Since then it has revolutionized laboratory study of T cells.

Commercial interests have, however, been stimulated by evidence that T cells can kill tumour cells and that interleukin-2 can stimulate T-cell killing of tumour cells both in the laboratory and within the body. This anticancer potential could transform the present annual \$1 million laboratory reagent business into a potential multi-million dollar pharmaceutical business. Interleukin-2 may also have potential in diseases such as leprosy in which the immune system is chronically suppressed.

The first essential step to exploitation is the production of a pure product by recombinant DNA technology. Today's description of the sequence and expression in monkey cells of cloned DNA encoding the molecule puts Dr Taniguchi's group narrowly ahead of the competitors and a short step from the goal of expression in bacteria.

Surprisingly, Dr Taniguchi's institute has not yet reached a commercial agreement on the rights to use the cloned molecule and is still discussing it with several companies. Biogen SA, which claims to be close to the stage announced by the Japanese, already has an agreement with Shionogi & Co. Ltd to carry out clinical trials in Japan once its interleukin-2 is available. Biogen's work on the molecule has been carried out by Dr Walter Fierz in Ghent (Belgium).

Sandoz, the pharmaceutical company, has contracted the Genetics Institute in Boston for work on the production of interleukin-2. The project is being carried out by a group under research director Robert Kamen in collaboration with Robert Gallo at the National Institutes of Health, Bethesda, Maryland and Dr Kendall Smith at Dartmouth College, New Hampshire. The group says it is close to having the sequence of the cloned product and is working on two separate sources of the molecule.

Other biotechnology companies in the field include Quidel, a new company in San Diego directed by David Katz, Immunex in Seattle, established largely to exploit the potential of interleukin-2 by Steve Gillis and Chris Henney, and Cetus in San Fran-

cisco. All are close on the heels of Dr Taniguchi's group.

The prospect of several competing producers of recombinant interleukin-2 may raise further problems for the patent authorities in establishing the criteria for claiming priority. One possibility is that the inevitable lengthy delay in reaching a deci-

sion would be sufficient, if interleukin-2 proves effective in clinical trials, to exploit the new molecule.

Interest in interleukin-2, however, does not stop at the molecule itself and work is under way to produce synthetic peptides that can substitute for it. Complementary approaches may follow from the recent production of a monoclonal antibody that recognizes the receptor for interleukin-2 (see *Nature*, 300, 267; 1982).

Nigel Williams

Anthropologist fired

Stanford plays its China card

Washington

THE ethics and obligations of Western anthropologists working in the People's Republic of China have been highlighted by the dramatic firing by Stanford University's anthropology department of a doctoral candidate, Steven Westley Mosher. The faculty's vote of 11 to 0 against Mosher is apparently unprecedented. Mosher says he will decide in coming weeks whether to sue Stanford for the PhD degree he says he has earned. "I will fight this thing to the end", he said. In the meantime, he is finishing a book, *Broken Earth*, about life in rural China. Stanford's reasons for the dismissal of Mosher as a doctoral candidate are not known, but are contained in a 47-page report by a faculty committee which neither Stanford nor Mosher will make public on the grounds that its contents could hurt "innocent parties".

The controversy revolves around bizarre tales of Mosher's conduct while he lived in an unnamed Chinese village on the Pearl River delta in 1979-80. His stay was sponsored by the respected Committee on Scholarly Communication with the People's Republic of China (CSCPRC).

Mosher's thesis for Stanford was about farming and fishing in Taiwan and he went to mainland China to study rural village life. While there, he became outraged by the practice of forced abortions for women who were seven, eight and nine months pregnant and what he alleged was officially sanctioned infanticide of girl babies in his commune.

Mosher says that in 1980 he informed Vice-Premier Chen Muhua, head of the Chinese birth control programme, about these abuses in the "naive" belief that the authorities would stop them. After leaving China, Mosher published an article about the cruelties of Chinese birth control practices in a popular Taiwan weekly. The article included photographs of women seven or eight months pregnant being prepared for abortions.

"The photographs taken by Mr Mosher during his story", wrote the Chinese authorities to Stanford, as part of the proceedings in the case, "exceeded the scope of his research topic (and) had nothing to do with any conceivable social science research topic". Mosher wrote in rebuttal:

"It is a measure of the success of my research that the Chinese Communists are so anxious to discredit it".

When complaints about Mosher's conduct in China reached members of the Stanford faculty, particularly the two anthropology faculty members who had originally endorsed the trip, the department began to investigate. The dissertation committee originally formed to review Mosher's Taiwan thesis disbanded, although Mosher says he has finished his dissertation. "I was three faculty signatures away from the PhD", he says.

Mosher's conduct while in China has also been the subject of much comment, not only by the Chinese, but by other

Victory for Yale

Washington

YALE University has won a round in the continuing fight between universities and the federal government over "effort reporting" by faculty receiving federal research grants.

Under an agreement reached earlier this month, Yale faculty members will no longer be required to complete detailed reports on their allocation of time between teaching, research and administrative duties. Instead, they will simply have to sign a statement verifying that direct charges made to a research grant represent work done on the project.

Indirect costs, or overhead, will be reimbursed at a fixed rate based on the experience of the past few years. Previously, the reimbursement of indirect costs was tied to the amount of time faculty reported spending on departmental administration. Stanford University has a similar arrangement providing for a fixed rate for indirect costs.

The universities have opposed detailed reporting requirements, repeatedly sought by the federal Office of Management and Budget (OMB), arguing that there are no clear lines separating a faculty member's various duties. Yale became a focus of this opposition when Professor Serge Lang of the mathematics department publicly refused to comply with OMB's requirements.

Stephen Budiansky

anthropologists who were in China as well and by his former wife, who is Chinese. His alleged misdoings range from financial irregularities to smuggling, and being an intelligence agent, which he denies. The head of the Stanford anthropology department, Clifford Barnett, says the secret faculty report does not even discuss the rumour that Mosher was an intelligence agent.

Stanford interviewed Mosher's ex-wife but is not telling Mosher her present whereabouts. Under Stanford's procedures Mosher was given a chance to see the "evidence" against him and to refute it.

The incident raises questions about how anthropologists should conduct themselves in states such as China unused to having Westerners observe ordinary life.

The Stanford faculty has issued a statement saying that his publication of the Taiwan weekly article did not amount to "misconduct". The statement does, however, say that the anthropologist should be "open and honest in accounting to others in the profession about his behaviour in the field". It says anthropologists should "pro-

tect the confidentiality" of those who provide them with information and "anticipate" the impact on the populations they study of dissemination of their research. Further, the statement mentions "protecting the good name" of the discipline by "abstaining from activities while in the field that are judged to be 'illegal' in areas where they work".

Other anthropologists are sympathetic to Mosher's outrage over Chinese birth control practices, accounts of which have filtered out of China in recent months. Thus Kenneth Pruitt, president of the Social Sciences Research Council, who is familiar with the case, says that Mosher had an obligation to let world public opinion know. "But he also had an obligation to put it in perspective." Pruitt acknowledged that other factors besides the birth control issue had been important in the Stanford faculty's decision.

Says Mosher, "I never had a course in ethics in anthropology when I was at Stanford. Now they give a course in it. And you can be sure that it is thanks to Steven Westley Mosher." **Deborah Shapley**

Fluoridation

Verdict awaited

JUDGEMENT has now been completed in the case, which may be the first of a series in Britain, to decide whether a regional council (in this case Strathclyde) may legally add fluoride to its domestic drinking water supply. Lord Jauncey, an eminent Scottish judge, has been considering the question since last July after a petition that took a record 205 court days to hear and costs more than £1 million. It is understood that the judgement is now being typed, which may itself take some time if the volume of evidence in the case is anything to go by.

In October 1978 an interim interdict was applied for in the Edinburgh Court of Session to prevent the regional council, as the water authority, from adding fluoride to the water supply of an old-age pensioner living in Glasgow, Mrs Catherine McColl. The hearing before Lord Jauncey started on 23 September 1980: Mrs McColl made a brief appearance at the start of the petition and has not been seen since. Subsequent witnesses included the well known US opponents of fluoridation, Dr John Yiamouyannis and Dr Jean Burk, who concentrated on their contention that fluoride can cause cancer (although sundry other ailments are listed in Mrs McColl's petition). The link with cancer comes almost entirely from a study by Dr Yiamouyannis published in the journal *Fluoride* of the International Society for Fluoride Research (the journal is not refereed and Dr Yiamouyannis is a co-editor).

For the water authority an impressive range of expert witnesses argued against the claims of the anti-fluoridation lobby, notably Sir Richard Doll, former Regius Professor of Medicine at the University of Oxford and Professor J.J. Murray of the University of Newcastle upon Tyne. These and others spent many days explaining methodological flaws in the Burk-Yiamouyannis evidence and studies purporting to show that fluoride is ineffective in preventing dental caries, and also presented contrary evidence of their own.

The water authority is required by law to provide "wholesome water", but Parliament has never attempted legislation on the fluoride question — although 10 per cent of the British population now drinks artificially fluoridated water. Strathclyde will be a test case for Britain. The courts in the United States have consistently found that water fluoridation is not a violation of individual or religious liberty, and more than half the states have ruled that any infringement of freedom to act by mandatory fluoridation is minimal, compared with its value as a preventive measure.

The health authorities in Scotland that first asked for fluoridation are resigned to a long wait. Whichever side loses will appeal, and the case is likely to end up in the House of Lords. **Tim Beardsley**

Oil pollution

North Sea still suffering

Brussels

IN spite of the bewildering number of international conventions that have been negotiated, the voices clamouring for urgent action to clean up the North Sea are growing stronger. At the last meeting of the European Community's environmental ministers at the end of February, the West German Government put forward its suggestions for another international conference on the problem. And at the European Parliament's March session, the European Commission was urged to create a North Sea pollution police force.

Next October, the Marpol treaty will come into force, adding yet another international agreement to the Oslo, Bonn, Paris and London treaties. Yet all the evidence available, and in particular the

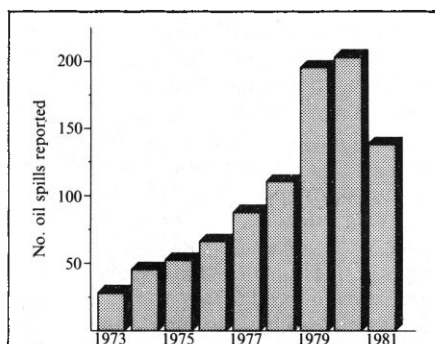
disturbing reports from Heligoland Bird Observatory, points to increases in the number of birds killed by oil pollution and the ineffectiveness of these treaties in controlling the discharges made by ships at sea.

The West German memorandum was well received by the other EEC countries. It stressed the need for intensive aerial surveillance of the North Sea to trace the origins of the oil spills, tougher fines for those caught polluting and improved technical pollution control at ports for oil and other installations. The goal of the proposed conference would be to find practical ways of compensating for the shortcomings of existing management policies.

The European Parliament proposed that ships carrying dangerous cargoes should have to carry a "black box" similar to that used in aircraft to provide an automatic recording of a ship's movements. But, as one Belgian parliamentarian pointed out, phosphates, chlorides and radioactive materials also need monitoring.

In October, environmentalists in the Netherlands will be trying to lay the blame on some of the worst offenders. An international water tribunal will present cases against major petrochemical companies. Independent jurists will be asked to judge the cases and the accused will be able to defend themselves.

Meanwhile, the *Amoco Cadiz* disaster of March 1978 is now coming to court in Chicago, where 76 Breton communities and the French Government are seeking compensation. In a comparable case in the United States, the damage claim was honoured to the tune of \$43,000 per tonne of spilled oil. **Jasper Becker**



Oil spills: The number of incidents reported between 1973 and 1981 in eastern Scotland, the Orkneys and the Shetland Islands (those parts of the British coastline which border onto the North Sea) and the North Sea offshore oilfields. Source: *Annual Reports of the Advisory Committee on Oil Pollution of the Sea 1973-1981*.

Yellow rain

Australian report goes public

Washington

FURTHER details of an Australian Government analysis that concluded that alleged "yellow rain" samples were fakes emerged last week as the new Labor Government agreed to release the report of the analysis (see *Nature*, 17 March, p.200).

The report, dated August 1982, had not been publicly circulated. Speculation has it that this was to avoid embarrassment to the United States, which has repeatedly claimed to have proof that yellow rain is a toxin weapon being used under Soviet supervision in South-East Asia.

The Australian analysis, performed by Dr Hugh Crone of the Department of Defence Support Materials Research Laboratory, concludes that "the samples examined are not toxic and in fact are composed of yellow pollen grains, probably with a small amount of binder". The pollen appears in the form of yellow spots adhering to leaves and pebbles, which, the report says, "were deliberately applied, either by a brush or by a spraying process". The report explicitly rules out the possibility that the spots are of natural origin.

The pollen in most of the samples matched that of the genus *Harpullia*, a group of rain forest trees that includes the tulipwood. The report notes that "one tree producing pollen from many flowers at one time would furnish enough material" to produce all the samples received by the Australians. Pollen from other tropical plants (*Sterculiaceae*, *Dilleniaceae*, *Cupaniaceae*, *Sapindaceae*) and from cereals (*Poaceae*) and the daisy family (*Compositae*) were also found in smaller quantities.

According to Phil Harrison of the Australian Embassy in Washington, the samples were turned over to the Australian Embassy in Thailand by a Laotian refugee in mid-April. Harrison said the Australians, who had earlier offered to analyse any material relating to "yellow rain", were informed by the American Embassy that the samples were available.

The report notes that because little systematic study has been made of pollen in South-East Asia, "it has not proved possible to identify grains to a level that could assist in locating the area of origin of the samples within South-East Asia". The bulk of the pollen, though, is probably derived from rainforest, the report says; the small amount of cereal and weed pollen present in some of the samples "could suggest that samples were taken from secondary forest trees close to cleared agricultural areas".

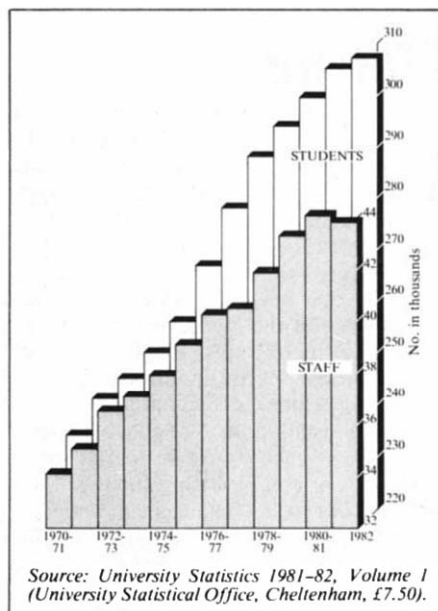
Unlike the US analyses, the Australians' did not specifically analyse for trichothecene mycotoxins, the fungal toxins that the United States maintains are the deliberate — and lethal — component of yellow rain. But the Australians did perform a simple

toxicological test (application to shaved rat skin), which the report says is sensitive to 500 ng of DAS or T-2, two of the chief mycotoxins. The results were all negative. "If we accept 500 ng as the upper limit, this cannot represent a militarily-effective residue", the report concludes.

The report also notes the presence of fungi, including *Fusaria*, on some of the samples. *Fusaria* are the natural source of mycotoxins.

Meanwhile, until more is known about the pollen that has reportedly been found in many, if not all, of the samples that the United States and other countries collected in Thailand, the only sure conclusion is, as the report states, that "since the [Australian] samples are obviously fakes, they convey no information at all as to the veracity or otherwise of the reports of chemical attacks". **Stephen Budiansky**

University staffing



THE total number of full-time students at university in the United Kingdom, both undergraduate and postgraduate, has risen by 31 per cent since 1970. This has been matched by a corresponding 29 per cent increase in the number of full-time staff, so that the ratio of students to staff has remained almost constant throughout Great Britain during this period, being 7.6 to 1 in England (standard deviation 0.13) and 7.5 in Wales and Scotland (s.d. 0.12 and 0.28 respectively), and has improved in Northern Ireland, falling from 10.4 in 1971-72 to 8.6 in 1980-81. However, as the number of students continues to rise, albeit at a rather slower rate, and the number of staff begins to fall, the reduction in the staff to student ratio in 1981-82 can only continue.

Melanie Kee

Sizewell B inquiry

Opposition falters

ALL is not going according to plan for the Sizewell B Appeal Fund, set up in January to support objectors at the marathon public inquiry now taking place at Snape in Suffolk.

The fund hopes to raise £0.5 million to finance groups at the inquiry opposing the plans of the Central Electricity Generating Board (CEGB) to spend £1,100 million on building Britain's first pressurized water reactor (PWR) at Sizewell. But the chairman of the fund's Appeal Committee, Mr Edward Irving, says that so far the fund "has not attracted the support we have hoped for".

The fund got off to a promising start when seven distinguished trustees were appointed to oversee applications for support within three weeks of the fund's foundation (*Nature* 10 February, p.457). The National and Local Government Officers Association, which has declared itself in favour of nuclear power, gave £1,000 to the fund in February, saying that it believed the donation would "assist towards the achievement of a fair and well-informed decision by the inspector". However, no other unions are known to have contributed, and approaches to industrial companies have so far failed to produce much positive response. The Appeal Committee may have to reconsider the fund's future unless more money arrives soon.

At the inquiry itself, CEGB and British Nuclear Fuels Limited have completed their initial presentation of evidence before the Inspector, Sir Frank Layfield QC. Objectors are now getting the first real chance to show their mettle with the cross-examination of Mr R. Priddle, the Department of Energy's only witness. The thrust of Mr Priddle's evidence was that it would be dangerous for the country to rely excessively on coal for its future energy requirements.

During cross-examination from Mr John Blake, vice-chairman of the Town and Country Planning Association, Mr Priddle acknowledged that the government had moved away from the 15-GW, 10-year programme of nuclear power station building that was planned in 1979. The number of new nuclear power stations that the government would like to see built in the next decade is now thought to be nearer four or five than the ten originally foreseen. Lower predictions for future electricity demand have inspired the shift: indeed, demand for electricity in the United Kingdom actually fell by 0.1 per cent last year. However, CEGB expects demand for electricity to increase to the end of the century even if total energy demand falls.

The importance of the nuclear power programme is now framed more in terms of its strategic contribution (diversifying

means of supply) and the savings made through the lower generating costs claimed for PWRs. This shift in government policy will be seized upon by the objectors, who this week include the Campaign for Nuclear Disarmament and the Stop Sizewell B Association.

The stress on diversity will be used by environmentalist to persuade the inspector that CEEB has given insufficient attention to the potential of renewable energy sources such as wave and wind power. The lower generating costs claimed for PWRs are also likely to be challenged, since CEEB's economic case appears to rely on future increases in the cost of coal. A recent CEEB study has revealed that so far electricity generated from nuclear power in the United Kingdom has been more expensive than electricity from coal.

The Nuclear Installations Inspectorate (NII) will take the stand in April to start their evidence on safety aspects of the pro-

posed design. It is clear that NII will not have completed its examination of CEEB's design by this time: on 1 March it detailed 80 outstanding issues on which it is seeking more information from CEEB. The inspector may consider an adjournment to enable NII to extend its presentation. NII has denied responsibility for delays in providing information on safety issues and blames CEEB for failing to provide NII with its Pre-Construction Safety Report until 10 months after the scheduled date. NII says "there can be no short cuts in carrying out the licensing process, and progress will depend on the timing and quality of the material provided".

It is to be hoped that Sir Frank Layfield has a hardy constitution: the inquiry is already expected to last at least until the end of the year. He is not unfamiliar with the subject, however, having represented some of the objectors at the Windscale inquiry in 1977.

Tim Beardsley

Soviet Jewry

Request for help with status

Jerusalem

MOSCOW'S community of scientific "refusniks", Jewish scholars barred from academic life after filing an application to emigrate to Israel, have issued a decalogue of "requests" for help to Western colleagues. Their message was delivered last week to the World Conference on Soviet Jewry here. This, the third such gathering since 1971, included a colloquium to discuss the worsening situation of Jewish scholars and scientists in the Soviet Union, whether or not they wish to emigrate.

Those concerned request in particular that all help tendered should be on the basis of their status as scientists, not as charity. Their statement says that invitations to foreign scientific conferences can be of particular help. Although an exit visa for the term of the conference will almost certainly be refused, such invitations serve as reminders to the authorities and may even lead to a reassessment of their status. The statement also asks that scientists invited to the Soviet Union should use the opportunity to get in touch with refusniks and that distinguished foreign scientists should put pressure on the Soviet Academy of Sciences. Scientists from international bodies such as CERN, the European centre for nuclear physics, with which the Soviet Union wishes to maintain good relations, may be especially influential.

Other suggestions in the statement are:

- Foreign scientists should attend the unofficial seminars organized by the refusniks, thus helping to avoid police interference.
- Western scientists are asked to help refusniks to get their papers published, particularly in seeing papers through the publication process.
- Refusniks need scientific journals, as they are excluded from academic libraries; in

the case of those sentenced to Siberian exile, the refusniks' journals should be sent to the next of kin.

The statement also asks that scientific journals should deal with the refusniks' plight, pointing out that articles dealing with the subject are omitted from the photocopied versions circulated by the central library service in the Soviet Union. They have heard, however, that negotiations are soon to begin for cover-to-cover translations of *Nature*, *Science*, *American*

Journal of Physics and other leading journals to be published in Russian, and they call on the respective editors to insist that no such excisions are made as a condition of agreeing to such an arrangement.

This last point illustrates the vacuum in which the refusniks live — *Nature*, for example, has heard nothing of such a project. Nor do the refusniks realize that while many scientists are prepared to visit the unofficial seminars as private individuals, it may not be possible for them professionally to have their names openly linked with them. The question of publication of refusniks' articles — as many participants in the scientists' colloquium reiterated — is a very painful one; much as they would like to help, the information lag almost inevitably means that the articles are outdated before they ever reach a Western editor's desk.

With exit visas for Jews down to a dozen or so a week, heavy policy pressure on refusnik seminars and increased pressure on those Jewish academics serving sentences in labour camp or exile, the participants in the scientists' seminar decided to set up an international coordinating secretariat to synchronize and render more telling protests by the scientific community.

In one respect, however, their efforts will no longer be needed. During the past two years, at least ten Jewish scientists have been deprived of their higher degrees for exhibiting an insufficient degree of Soviet patriotism. This practice, the refusniks say, has now ceased — the authorities have abandoned it as a tactical error, apparently as a result of Western condemnation.

Vera Rich

Acting ambassadors

SCIENTISTS and scholars tend to be regarded by the Soviet authorities as unofficial ambassadors for the Soviet system. For this reason they prefer to send to an international conference some staid and elderly personage who can be trusted to behave with decorum, rather than some brilliant young researcher who might come out with an indiscreet criticism of the Soviet Union.

This tacitly understood role for the Soviet scientist was more clearly defined recently by Academician Georgiy Skryabin, a principal academic secretary of the Soviet Academy of Sciences. Addressing the annual general meeting of the academy this month he praised the "great work" done by scientists abroad at conferences and study trips in "explaining the resolutions of the twenty-sixth party congress, and the internal and external policy of the Soviet Union".

The main public relations service the scientists have performed, Dr Skryabin noted, related to Soviet proposals for peace and security. They had spoken up in international forums against the threat of

nuclear catastrophe.

Nothing was said at this meeting, however, of the possibility (widely discussed in Moscow just before Brezhnev's death) of a Soviet initiative to form a Soviet version of an international organization of scientists against nuclear war parallel to the existing doctors' organization. Indeed, Academician E. P. Velikhov, who had been widely tipped to head the putative organization, and who recently gave a major interview on the dangers of nuclear war on Moscow television, confined his remarks at the meeting to the importance of organizing the academy's research on automation and computer science. The peace issue was well to the fore, though, with the awarding of the 1982 Lomonosov medals, the highest award of the academy, presented annually to one Soviet and one foreign scientist. This year, the foreign recipient was Professor Dorothy Hodgkin of the University of Oxford, whose campaigning for peace was especially commended by Academician Yuriy A. Ovchinnikov in his keynote speech "Marxism and Scientific-Technological Progress".

Vera Rich

Laboratory animals

European convention in sight?

EVERY six months or so since 1978 an *ad hoc* committee set up by the 21-nation Council of Europe has met in Strasbourg, France, to perfect a "Convention for the protection of animals used for experimental purposes". The committee's deliberations are at last nearing an end, and by the end of the next meeting, on 26-29 April, the convention may be nearly ready.

In the same week that the committee meets, the democratically elected European Parliament will also be airing the issue. Last December the parliament took evidence from animal welfare groups and others at a public hearing in Strasbourg. This evidence is now being assessed and the draft version of the convention is on the agenda for the parliament's next sitting at the end of April. Eventually the parliament will make recommendations to the EEC Council of Ministers.

In Britain there has been steady pressure for new legislation to regulate the use of animals in laboratories, at present licensed under the Cruelty to Animals Act of 1876. In the past few years two private members' bills have floundered, one in the House of Commons and one in the House of Lords, and one reason for the absence of government support that might have seen one of the bills through to the statute book was the prospect of a European Convention on which to base the national law.

The five British committee members (the "academic" representatives are Professor Alice Steele-Bodger and Dr Charles Coid of the Home Office advisory committee on animal experimentation) must make their contribution to the final wording of the convention against the background of an increasingly aggressive campaign against the use of animals in research on any terms. Recent action by animal welfare groups has included break-ins to release animals from laboratories, and last weekend demonstrators gained considerable publicity by chaining themselves to the railings outside the Prime Minister's residence in Downing



Street (result shown below).

So whilst few scientists are likely to claim that an act passed in 1876 is comprehensive enough to regulate the wide range of uses of animals today, updating the law has intrinsic risks when the major pressure for reform comes from groups that see it as only a first step towards banning the use of animals in research altogether.

The Royal Society for the Prevention of Cruelty to Animals (RSPCA) claims to have a more realistic view than the extremists. At the parliamentary public hearing last December, the RSPCA presented evidence on behalf of the Eurogroup for Animal Welfare aimed at safeguarding the welfare of animals in laboratories. The RSPCA's case not surprisingly concentrates on the pain that animals might suffer, and the "pain condition" is a subject that has contributed to the delays on the *ad hoc* committee. The draft convention states that if an animal suffers severe and enduring pain the experiment should be stopped and the animal killed, and the argument has been about whether there is a need for exemptions to the clause.

One of the convention's basic principles is that animals used in experiments should



be specifically bred for the purpose. This point is considered of vital importance by animal welfare groups, although researchers who work on stray, feral or farm animals, many destined to be killed anyway, claim that such a restriction would needlessly increase the number of animals bred for research, hardly in line with the goal of those who would protect animals.

And another of the requirements in early drafts of the convention might lead to animated discussions: "An animal shall not be used in a procedure if another satisfactory method is reasonably and practically available". That should add fire to the usually inconclusive arguments between tissue culture protagonists and others over how well an *in vitro* system can or cannot mimic the animal. **Charles Wenz**

UK education

A central role

THE British Government appears to be heading for a conflict with school-teachers as well as with their employers, the local education authorities which are statutorily responsible for running primary and secondary schools. By convention and on one reading of the 1944 Education Act, the content of what is taught in schools is the responsibility of local authorities alone, and the once-formidable corps of Her Majesty's Inspectors of Schools have become largely advisory.

Since 1976, however, British governments have become more adventurous, offering advice on the ideal content of a "core curriculum" for example. Last week, however, Sir Keith Joseph, Secretary of State for Education and Science, went further than any of his predecessors by issuing detailed instructions on physics examinations to be administered to students at 16-plus by the newly formed joint council of examinations boards.

The specificity of Sir Keith's opinion will give the most direct offence. While the examinations council had proposed that experimental skills in physics might be tested by means of written examination questions, Sir Keith now says that only a practical test will suffice to measure "the combination of intellectual and manipulative skills" he seeks to encourage.

Sir Keith's letter also takes issue with the council's view that electronics should not

be a part of its proposed core of secondary-school physics, asking that its proposals should be modified in such a way as to ensure that all students should learn something about electronic devices. On the other hand, Sir Keith suggests that the examination proposals should "omit all reference" to the "social and economic issues" arising from scientific developments on the grounds that students would then be distracted from the main objectives of science teaching and because it would be difficult to "avoid tendentiousness in the teaching of science subjects".

Reaction to these developments is, for the time being, stunned silence, but it will obviously now be harder for the new examinations council effectively to discharge the job for which it was created — simplifying and coordinating the jungle of 16-plus examinations at present administered by a score of different examinations boards.

The impending row about examinations will be further inflamed by the appearance this week of a Government White Paper (*Teaching Quality*, Cmnd 8836, HMSO, £3.40) in which Sir Keith Joseph, assisted by his opposite number for Wales, outlines a series of measures for the improvement of the quality of teachers. These range from a proposal that the selection of intending teachers, and the criteria for approving the training courses followed, should be made more rigorous to a requirement that those who train them should have recent practical experience of teaching in a school and a suggestion that good teachers might be paid extra. [1]

Israeli research manpower

Shortage in prospect

Rehovot

THE need for a substantial increase in the number of academically trained people to work in research and development is widely accepted. The National Council for Research and Development has urged that Israel will need 86,700 such people in 1995, compared with 34,800 in 1974 — an increase of 150 per cent. But there is vigorous disagreement about the steps that need to be taken to ensure that the well-trained researchers actually materialize.

Thus, Professor Josef Singer, president of the Haifa Technion, claims that there will not be enough such people because the institutions of higher learning are being starved of funds. He points out, for example, that there was a 26 per cent increase in students at these institutions between 1974 and 1982, while funds increased by only 18 per cent. If spiralling costs are taken into consideration, he adds, the picture is grimmer still.

The Technion is adversely affected by yet a third factor, the rapid development of science-based industry. High-technology firms now offer much higher salaries than the Technion, which is consequently suffering from a lack of instructors in such fields as computers and electronics. The result is that it cannot train as many students as its physical facilities permit.

Singer's views were, however, challenged by Professor Derek de Solla Price of Yale University, who was in Israel last week for a conference on microcomputers. In the 1960s Price developed a method for measuring scientific manpower in various countries based on the total number of researchers who had papers published in major professional journals, and concluded that Israel then had five times as many scientists as would be expected for its population and gross national product.

Price insists that "the situation is no different today; Israel still possesses an enormous reservoir of trained people, something for which she has every reason to be grateful because her scientists and technicians more than compensate for her lack of oil and minerals". But in Price's view, they must be fruitfully employed, which means that more should be engaged in industrial research and development. Changes should begin, moreover, in Israeli high schools, where Price calls for better experimental training and equipment because "what Israel requires are people with brains in their fingertips".

Professor Haim Aviv, who has taken leave from the Weizmann Institute of Science to establish the company called Bio-Technology General in Rehovot's Kiryat Weizmann Industrial Park, agrees that the country has plenty of researchers, but bemoans the lack of engineers for development work in the rapidly expanding field of biotechnology. Elsewhere, he

says, newly established companies have been able "to steal" the men it required from pharmaceutical companies, a reservoir of talent that does not exist in Israel.

This problem might not have arisen, Aviv adds, had Israel's National Council for Research and Development dealt with long-range planning, as it should have done. But it lacked adequate power, a situation he hopes will change now that Israel has a Ministry of Science and the council has become part of it.

Yet even the establishment of a Ministry of Science does not guarantee a uniform approach to the utilization of research and development manpower. Israel's Science Minister, Professor Yuval Ne'eman, is a supporter of major projects such as the building of a nuclear reactor (if none can be purchased overseas), the production of a new Israeli warplane (the Lavie) and the construction of a Mediterranean-Dead Sea Canal (for power production and other purposes).

In contrast, Professor Arie Lavie, chief scientist of the Ministry of Industry and Trade, puts more emphasis on a large number of smaller enterprises. He points out that smallish companies have made it possible for Israel to develop many high-technology products which already account for more than 30 per cent of its total industrial exports. In addition, Lavie sees less risk in this approach because the inevitable shaking-out process will be less painful if a few small companies fail than if one or two major enterprises prove misguided.

Israeli planners must also take into consideration the unpredictability of immigration and emigration. They had no way of knowing, for example, that the country would receive 6,000 Russian-trained engineers or that a roughly equal number of Israeli-trained engineers would emigrate (although several hundred of them have recently returned, due to the economic downturn overseas and the development of high-technology industry here).

Military needs must certainly be met as well. Retired General Nati Sharon, who headed the planning branch of the Israel Defence Forces before taking up his present position with Tadiran (a major Israeli electronics company), recently warned that the army would soon lack the trained manpower needed to deal with the sophisticated equipment that it is now absorbing.

So far Israel has been able to meet all its needs for scientists, engineers and technicians by simply exploiting the manpower that happens to be available as the result of unplanned and uncoordinated developments. But, this may no longer be possible in an era when high-technology industries have become the fastest growing and probably the most important sector of her economy.

Nechemia Meyers

Falkland Islands

Salmon plan grounded

A PROJECT designed to produce a salmon farming industry in the Falkland Islands from the islands' two main resources — sheep and the sea — has run into difficulties. A feasibility study carried out by a group of British researchers headed by Dr James Muir of the Institute of Aquaculture, University of Stirling, has found that whilst the Falklands are well suited to salmon production, the only accessible markets for the proposed smoked salmon are several thousand miles away and already support a cut-throat salmon business. But the project supervisor, and director of the Institute of Aquaculture, Professor R. J. Roberts is optimistic.

There are two possible ways of raising salmon — farming and ranching. Ranching relies on the ability of salmon to "imprint" on the water in which they mature. Young salmon are released into a seawater loch, mature and move out into the open water. The salmon then return to their

*Never in Stanley?*

release-point to spawn, so a supply of fully-grown fish is obtained without an input of foodstuff. However, this economically sound system has run into biological problems and for the one successful salmon ranch currently being operated in Chile there have been dozens of failures.

It is for this reason that salmon farming is favoured by Professor Roberts. The aim is to feed the salmon on sheep meat, which is abundant in the Falklands. Sheep are raised for their wool, which is tough and used for carpets, and their meat is discarded. Tests at the British Tropical Products Institute have shown that salmon can be raised successfully on sheep meat.

The salmon farming project has been in the pipeline for several years. It would be ironic if the Falklands conflict, which has contributed to sudden availability of supportive funds, should also cause its downfall by taking away the most accessible markets in South America.

Melanie Kee

Mauna Kea best for new telescope?

SIR — A recent leading article in *Nature* (17 February, p.561) exposes the bad choice of original site for the Isaac Newton Telescope and this prompts me to draw attention to a likely mistake which could be made in locating the proposed UK millimetre-wave telescope. Mauna Kea in Hawaii appears to be the site now favoured but until the real properties of this site and its competitors are sufficiently investigated this choice should be questioned.

Millimetre waves are appreciably absorbed by water vapour and high altitude sites are chosen to minimize water amounts. However, measurements at several high sites, including Mauna Kea, have shown that the actual absorption can be higher than is predicted from the known amounts of water. While there is general agreement on the existence of an excess absorption, the nature and cause of it are controversial. But there is little doubt that the additional atmospheric phenomenon can, in some conditions, become the dominant factor in reducing atmospheric transparency to millimetre waves and this would be a serious impediment to would-be observers.

The UK millimetre-wave telescope project was started in the late 1960s following the discovery of a galactic carbon monoxide. I recall being asked by J.F. Hosie, then in charge of astronomy at the Science Research Council, if, when I joined the council's Appleton Laboratory in 1972, I would be willing to have a major involvement in the project because funding had been agreed and delay in building was giving him concern. In the event the then director of Appleton Laboratory, the late J. Saxton, took personal charge of the project and I was not involved. I did, however, persist in stressing the need for site evaluation and a 3-week measurement programme at Mauna Kea was agreed. The findings, reported in *Nature*¹, confirmed that Mauna Kea could be afflicted by excess absorption. They also gave indications that the source of the extra absorption was localized in the vicinity of the observatory.

I became convinced that there was a strong scientific case for investigating as an alternative the properties of desert sites at medium altitude, notably Kitt Peak. The succeeding director of Appleton Laboratory, F.H. Horner, refused to support this and comparison of maritime and desert sites has yet to be done. Failure to provide a rational basis of choice for siting a multi-million pound installation seems to be irresponsible. Whether the telescope should be built at all 12 years or more after it was conceived is questionable, but if the vested interests in it prevail, the taxpayer has a right to expect from the scientific community that all reasonable steps have been taken to prevent a second fiasco in locating a major astronomical facility. There can surely no longer be any hurry to

complete the millimetre-wave telescope — it still has 10 years to go to match the time taken to complete the Isaac Newton Telescope.

H.A. GEBBIE

Imperial College,
London SW7, UK

I. Moffatt, P., Bohlander, R.A., McCrae, W.R. & Gebbie, H.A. *Nature* 269, 676-677 (1977).

No delay on LEP

SIR — We were greatly surprised, at the French Ministry of Research and Industry, by your news item "Delay at Geneva", in the 10 February issue of *Nature* (p.459).

Talking about LEP, you state that "instead of final approval coming from the French constitutional court in the spring, as expected, it might not arrive before the end of the year". This assertion is completely wrong: the Conseil d'Etat (not the Constitutional Court, which has nothing to do with such problems) will be in possession of all documents shortly and approval is expected within a few weeks, using a procedure of emergency. The decree of "Déclaration d'Utilité Publique" will be signed immediately after by the prime minister.

Of course, the installation of such a large device as LEP may cause some trouble — and also bring some advantages — to the inhabitants of the Pays de Gex. This is why, in our country, very serious procedures have to be gone through before approval can be given, the most important ones being the "Etude d'Impact" and the "Enquête publique". So far all these procedures have taken place according to the law, without significant departure from the time limits predicted a year ago. If the construction of LEP is subject to delay, it will not be because of the approval procedure.

P. PETIAU

Ministère de la Recherche et de
l'Industrie, Paris, France

No to astrology

SIR — In seeking to defend astrology¹ Richard A. Batchelor makes several errors.

He offers spurious accuracy to the pseudoscience by noting that astrology is "based upon a unique permutation of 10 planet positions $\times$ 12 zodiac signs $\times$ 12 'houses' calculated from birth time, birth date and place". He does not even mention the myriad disagreements between astrologers about the nature and import of the various houses and planetary aspects, nor that the most commonly used systems deprive whole nations of any astrological destiny by virtue of having no houses in which to place their guiding planets².

He further asserts that astrology's credibility lies in "empirical evidence". The most thorough review of which I am aware³ indicates that the vast majority of actual studies find no evidence whatever in

favour of astrology. J.H. Nelson's oft-cited work on planetary alignments and radio-wave propagation⁴ is mentioned favourably. But the demonstrated correlation is the product of a simple statistical error^{5,6}, a discovery that has been widely published⁷.

Batchelor's jibe about the sixteenth century church reminds me of Stephen Jay Gould's remark⁸, "a man does not attain the status of Galileo merely because he is persecuted; he must also be right".

JEREMY CHERFAS

Animal Behaviour Research Group,
Oxford University, UK

1. Batchelor, R.A. *Nature* 302, 10 (1983).
2. Culver, R.B. & Ianna, P.A. *The Gemini Syndrome* (Pachart, Tucson, Arizona, 1980).
3. Dean, G. & Mather, A. *Recent Advances in Natal Astrology* (The Astrological Association, Bromley, 1977).
4. Nelson, J.H. *RCA Rev.*, March, 26-34 (1951).
5. Ianna, P.A. & Margolin, C.J. *The Skeptical Inquirer* 6(1), 32-39 (1981).
6. Meus, J. *The Skeptical Inquirer* 6(4), 30-33 (1982).
7. Chérfas, J. *The Guardian*, 3 November (1981).
8. Gould, S.J. *Ever Since Darwin*, 154 (Norton, New York, 1977).

Tackling homology

SIR — In entitling their article "Functionally homologous cell cycle control genes in budding and fission yeast", Beach *et al.* (*Nature* 23/30 December 1982, p.706) erred in their usage of the term homology. Such usage reflects a common confusion about the relationship between function and homology. Their finding that the *cdc 2* gene of *Schizosaccharomyces pombe* and the *cdc 28* gene of *S. cerevisiae* perform similar functions indicates only that they serve functionally *analogous* roles. The authors' further contention that the genes and molecular mechanisms controlling the initiation of mitosis in *S. pombe* and *S. cerevisiae* were highly conservative through evolution, and thus homologous, is an entirely different issue. In failing to keep their study of function separate from their speculation on phylogenetic similarities, the authors used the unpalatable phrase "functionally homologous".

If similarity of function is accepted as a criterion for the determination of homology, the term will cease to be of use to biology as it will no longer differentiate between characters inherited from a common ancestor and those which arise from convergent evolution. Furthermore, to describe the relationship between the *cdc 2* gene and the *cdc 28* gene as one which lacks "sequence homology" but is "functionally homologous" implies that these are different "kinds" of homology. Qualifying homology in such a fashion is like describing the wings of all the flying creatures in the animal kingdom as "functional homologues". In some instances they would be homologous, in others not; but function would never have anything to do with it.

THOMAS E. HUGHES

Department of Anatomy,
Duke University, North Carolina, USA

Do laboratory animals have rights?

The months ahead threaten more wrangling about the status of laboratory animals.

The long-promised European Convention could be a useful framework.

THE problem of laboratory animals, always with us, is likely to become more obvious in the months ahead as the Council of Europe nears the end of its attempt to negotiate an acceptable international convention on the subject (see page 284) and as the body of opinion that calls itself the animal liberation movement gathers strength, as seems likely. In Britain, the issue was made more pointed last week when Professor Peter Singer of the Australian National University delivered a television lecture as part of a kind of celebration of the British Broadcasting Corporation's "Horizon" programmes (often rebroadcast as part of the Nova series in the United States). Singer, a philosopher by trade, has won a reputation among those who break into laboratories and steal the animals they find there, and now in Britain send letter bombs through the post, as the man who makes their vandalism respectable — an allegiance that Singer himself refuses.

Singer's argument is interesting and provocative, and deserves attention even by those who disagree with it, or with its conclusions. The argument has three components. First, Singer says, the notion that animals may properly be used as human beings calculate will best serve human interests is akin to older notions that barbarians or aborigines could be killed (or made into slaves) on the grounds that they were lesser beings, and is just as ethically improper. "Speciesism" is Singer's word. Second, to justify the mistreatment of animals on the grounds that interspecific competition is unavoidable in evolutionary terms is unjustifiable and at odds with moral progress (if not practice) in recent years. Third, it does not however follow that all animal species are equal, but rather that in the design of experiments in which animals are used, those responsible should put themselves in the animal's shoes, asking themselves how they would welcome what is planned, in which case (Singer rightly says) a relevant consideration is that as much as possible should be learned by observation of the behaviour of the non-human species so as to estimate its capacity for suffering.

The argument is consistent enough. It does not of course imply that animal experiments cannot be carried out, but merely that there would be fewer of them. Singer, fair play, reinforces his conclusion with several provocative examples. Thus

on the speciesism question, if the line is drawn between those with developed mental faculties and those without, why not use newborn children rather than dogs? Why not in any case use brain-dead people as experimental animals? And if chimpanzees are the only animal models of hepatitis B infection, but are an endangered species, why not ask for human volunteers instead?

Less radical arguments leading to very similar conclusions are also, of course, familiar. Thus even within the framework of what Singer calls speciesism, it is possible to argue that the practice of cruelty to animals debases human beings and runs the risk that familiarity with animal cruelty will engender cruelty to people. More to the point, wantonly to cause pain to animals is painful for those who witness these procedures, who learn about them or suspect that they may be going on. For what the animal liberation movement does not appreciate or accept is that even those who work in laboratories with laboratory animals find many of the procedures personally painful.

Largely on such grounds, most governments have now found it necessary to regulate the use of experimental animals by legislation. In the standard model, those who use laboratory animals must be registered, even licensed, and there are usually regulations requiring that animals should not be operated on without anaesthetics, that animal experiments should not be carried out in public and that animals should not be allowed to recover from damaging operations but should be killed instead. Needless pain is to be avoided. And then there are usually procedures by which exceptions to these rules may be permitted. The practical position, in other words, is not very different from Singer's.

So why not aim at legislation that will represent a reasonable compromise between the animal liberation movement and laboratory scientists, persuading both groups that since their chief objectives will be satisfied by the draft of the European Convention now circulating, there is no reason why they should not settle for an agreed prescription? For laboratory scientists, there are three snags: the convention would (and should) cover animals used for the routine production of antisera, would require that laboratory animals should always be bred especially for the purpose intended and is imprecise

about the circumstances in which exception from the strict rules will be allowed (and by whom). But should that hold up the signature of the European Convention, now so long put off that its non-existence must be offensive to those who want something done?

The danger is that further delay will play into the hands of people like Mr Singer. Singer's implied assertion that there is no sharp line to be drawn between people and other species of animals overlooks several simple truths. Populations of domesticated animals are controlled both as to number and to kind by people, not by the species concerned. While there may be no immediate interspecific conflict between people and other species now that smallpox has been eradicated, and although a practical way of eliminating *Anopheles* mosquitos would probably still be welcomed even though there is a vaccine round the corner, who can be sure that the survival of people this far has not been made possible by past interspecific conflicts? And is it, in any case, possible to think that the modern medicines for which the developed world is clamouring can be developed safely on the basis of clinical trials rather than with laboratory tests with animals? People might volunteer to be subjects in a trial of hepatitis B vaccines out of ignorance, but would hardly rush to help test the efficacy of some new analgesic.

The animal liberation movement is, for these reasons, doomed to failure but certain to cause trouble in the meantime. The scientific community would do well to answer positively, and not simply by means of the familiar case for research with laboratory animals that usually there is no alternative method.

Why not, instead, answer Singer's paradoxes with others? If the safety of some new medicine has not been assured by laboratory tests with animals, why not invite human volunteers and pay them a fee for their time and trouble? (It is not so long since the practice of offering prisoners in gaol a respite from their sentences was stopped by public demand, and rightly.) Or why should not the manufacturers of some promising new drug be allowed to advertise in the newspapers for those likely to benefit who are also prepared to take the drug untested? The answer to both questions is that governments, societies and people would not allow such things to happen. Mr Singer would no doubt complain of a resurgence of speciesism, but to no effect.

Third word development

Intensifying tropical agriculture: the Indonesian experience

from Gordon R. Conway and David S. McCauley

TEN years ago there were gloomy forecasts of widespread famine in many Third World countries as population increases were predicted to outstrip agricultural production. Indonesia, with a very high population growth rate and over 75 million people crowded onto the island of Java alone, seemed a prime candidate for Malthusian disaster. Yet, today, Indonesia has achieved near self-sufficiency in rice, its major staple food crop. Rice production, currently standing at over 22 million tons per year, has risen by an average annual rate of about 4.5 per cent over the past decade — well in excess of the 2.3 per cent population growth rate^{1,2}. This is a remarkable and praiseworthy achievement. Inevitably, however, the speed and scale of this development has brought with it a variety of problems — agronomic, ecological, social and economic. These were recently assessed at a small informal workshop held near Jakarta, under the auspices of the Ford Foundation.

About a quarter of the increased rice production has come from extending rice cultivation to marginal lands, in particular the tidal swamps of Sumatra and Kalimantan. Most, though, has come from doubling and trebling yields on existing paddy land by the introduction of new short-stalked fertilizer-responsive rice varieties (developed at the International Rice Research Institute (IRRI) in the Philippines), the rehabilitation of irrigation systems and the provision of subsidized inputs and credit to farmers.

Severe setbacks were experienced in the mid-1970s when recurrent outbreaks of brown planthopper (*Nilaparvata lugens*) devastated crops. Losses in 1977 were over 2 million tons — enough to have fed six million people for a year³. The outbreaks were brought under control through the introduction of a new resistant variety, IR36. Its use was vigorously encouraged and the growing of traditional varieties in the lowlands was prohibited. By 1980, IR36 had come to occupy well over one-half of Java's lowland paddy area. It matures in around 105 days and is non-photosensitive, so that double cropping became feasible even in most rainfed conditions. With year-round irrigation, either continuous rice cropping or a double rice crop followed by sweet potatoes, groundnuts, or soybeans was possible. Crop production again began to rise dramatically.

The brown planthoppers were encouraged by the widespread planting of susceptible varieties, continuous and overlapping cultivation, increased application of nitrogenous fertilizers and the destruction

of the insect's natural enemies by non-selective insecticides⁴. Massive increases in the use of pesticides has also resulted in many other serious side effects, including, it seems, reductions in the traditional ricefield harvest of fish — normally a major source of animal protein for the rural population.

Control of brown planthopper requires an integrated approach involving an appropriate mixture of resistant cultivars, selective insecticides and cultural measures. Breaks in the planting schedule followed by synchronous planting can help reduce the size and severity of the outbreaks.

In the meantime, the perils of putting all of one's eggs in one basket have been demonstrated — IR36 has proved susceptible to tungro virus (carried by the green leafhopper, *Nephotettix cincticeps*). In 1981 serious losses were experienced in Bali, Central Java and East Java. In response, the latest IRRI varieties (IR50, 52 and 54) were introduced, and following a crash programme of seed increase several hundred tons were distributed for the 1981–1982 crop season.

Constantly relying on the rice breeders' staying just one step ahead of the latest pest or disease outbreak is clearly a risky business. One solution would be a national breeding programme making greater use of indigenous germ plasm, so producing a broader spectrum of high-yielding varieties suited to a wide range of socio-economic and environmental conditions.

The recent pest and disease problems help to emphasize that agricultural development is not solely a question of increased productivity. Equal weight should be given to questions of stability (the year-to-year fluctuations in productivity resulting from environmental variability), sustainability (the ability to withstand repeated stress or a major perturbation)

and equitability (the distribution of output among the producers and consumers).

Research⁵⁻⁸ on the equitability of the new rice technology has largely focused on identifying changes in labour use. For example, before the introduction of the new rice varieties, rice was largely harvested plant-to-plant by women using the small *ani-ani* knife. The harvests were open to all community members — payment being a fixed proportion of the rice they harvested. As increasing numbers of landless labourers sought work, farmers turned to new harvesting arrangements, and the short varieties, which mature almost simultaneously, lent themselves to the use of the sickle — a tool which has come to be used primarily by men who work for wages in contracted harvesting teams.

There have also been changes in the equitability of post-harvest processing. In particular, mechanical hullers have largely replaced the hand-milling of rice by women, who traditionally received a fraction of the milled rice⁹. Even new community-level management techniques such as synchronous planting, which can increase water-use efficiency and pest-control, may adversely affect equitability. With synchronous planting, for example, the poorest farmers can no longer delay their own planting in order to supplement their incomes through wage labour on the larger farms. In general, the new technologies have contributed to increased incomes at all levels of the village rice economy, but this has been accompanied by a widening of the income and asset (including land) gap between rich and poor. Malnutrition of various types persists even in agriculturally productive villages. There are clearly serious trade-offs between increased equity and growth in rice productivity¹⁰.

Despite these problems, a recent survey¹¹ of 25 villages on Java indicates that the real wages of hired agricultural labourers have begun to rise and, more important, there has been a dramatic increase in off-farm employment opportunities both on and off Java. The question now is: how sustainable is this recent growth in agricultural productivity and relative rural prosperity? Further gains from increased use of chemical fertilizers, the lowland adoption of improved rice varieties and the rehabilitation of large irrigation systems seem unlikely. What remains is the considerably more difficult task of developing sustainable cropping systems for rain-fed agriculture, in particular in the uplands where serious soil erosion threatens both upland and lowland development. Difficulties with soil structure and fertility may also arise with increasing intensity of paddy culture.

At the same time, the government's ability and willingness to commit a large proportion of its resources to agricultural development efforts is becoming increasingly uncertain. Indonesia's oil and gas boom appears to be over, and this may reduce both the opportunities for off-farm

1. Agency for Agricultural Research and Development *Five Years of Agricultural Research and Development for Indonesia: 1976-1980* (Bogor, Gaya Teknik, 1981).
2. Central Bureau of Statistics *Monthly Statistical Bulletin*, 186 (May 1981).
3. Mochida, O. *Indones. agric. Res. Dev. J.* 1, 4 (1979).
4. *Brown Planthopper: Threat to Rice Production in Asia* (IRRI, Los Banos, Philippines, 1979).
5. Collier, W.L., Wiradi, G. & Soentoro *Bull. Indones. Econ. Stud.* 9, 36 (1973).
6. Collier, W.L., Soentoro, Wiradi, G. & Makali *Food Res. Inst. Stud.* 13, 169.
7. Hayami, Y. & Hafid, A. *Bull. Indones. Econ. Stud.* 15, 94 (1979).
8. Collier, W.L. *Development and Change* 8, 351 (1977).
9. Timmer, C.P. *Bull. Indones. Econ. Stud.* 9, 25 (1973).
10. Mubyarto, Sajogyo & Tjondronegoro, S. in *Growth and Equity in Indonesian Agricultural Development* (ed. Mubyarto) 215 (Yayasan Agro-Ekonomika, Jakarta, 1982).
11. Collier, W.L. et al. *Agroeconomic Survey, Bogor, Occasional Pap.* 6 (1982).

employment through decreased government spending and the extent to which the government can continue to hold down food prices, subsidize production inputs and support massive agricultural development programmes.

Equally worrying is Indonesia's ability to weather the inevitable bad years produced by new pest and disease outbreaks or poor rainfall. The traditional village welfare institutions based on the sharing of village product and labour are disappearing. The communal village rice stores are now often empty — replaced by more centralized government stores. It remains to be seen whether in the bad years the govern-

ment can act as efficiently as have the village-based institutions.

The national rice policy of the past decade has been highly successful, but now, with self-sufficiency at hand, the need in the next five-year plan, currently being drafted, is for a broader-based national food strategy. The variety of policy, technical and social questions which this raises are just beginning to be addressed.

Gordon R. Conway is Professor of Environmental Technology at Imperial College, London SW7 1LU and David S. McCauley is a resource economist with the Ford Foundation, Jakarta, Indonesia.

Texas symposium

Astronomy: relativity's relative?

from Virginia Trimble

THE uneasy marriage between general relativity and astronomy, solemnized by the biennial series of Texas Symposia on Relativistic Astrophysics, has endured another two years*, although the parties concerned continue to regard many of each others' concerns with tolerant puzzlement.

The relativists, who were rather badly outnumbered on the programme, have been worrying about such issues as the logical structure of general relativity, ways of improving numerical calculations within it and possible approaches to a quantum theory of gravity that might, someday, be unified with gauge theories of the other forces. S. Deser (Brandeis University) reported progress towards demonstrating that energy, properly defined, is always positive within general relativity, so that real systems are guaranteed to have ground states and stability. The proof is essentially complete both locally and globally for cases with zero cosmological constant, Λ , and for global-positive Λ solutions. The case with negative Λ is still not quite fully understood, but the speaker expressed confidence that it would all turn out all right.

T. Nakamura (Kyoto University) discussed results of numerical solutions for the collapse (in two dimensions) of a rapidly rotating star, in which the matter is properly included but gravitational radiation neglected. Considering the small number of grid points used, the results come impressively close to what we think should be the 'right' answer. Thus, for several different patterns of angular momentum distribution, most of the material collapses only when the ratio of angular momentum to mass is small enough for a stable Kerr black hole eventually to result. Larger ratios always lead to much of the matter being ejected, although the detailed ejection pattern is probably an artefact of the grid pattern, numerical viscosity and so on.

Finally, P. Van Nieuwenhuizen (SUNY,

Stony Brook) pointed out some of the pros and cons of supergravity, the extension of general relativity which adds to its gravitational field carried by gravitons a second field, carried by fermions (spin = 3/2) called gravitinos. On the plus side, it has now been demonstrated that supergravity (unlike general relativity) is renormalizable at least for all Feynman diagrams with one or two loops, though there are still some uncertainties at higher levels. A negative aspect is that it is not clear that even the largest ($N=8$) supersymmetry has enough freedom to include both non-zero cosmological constant (vacuum energy) and all the kinds of fermion field we know exist. The speaker expressed guarded optimism that further significant progress could be expected, perhaps along the lines of multi-dimensional space.

The astronomers, meanwhile, remain enamoured of telescopes and the objects they reveal. A half-day session addressed large ground-based optical telescopes for the late 1980s and beyond. The University of Texas 300-inch (7.6 m) project, discussed by H. Smith (University of Texas), is the furthest along in funding. It will probably be a thin monolith, with active support and/or honeycomb-style strengthening. The University of California TMT (Ten Meter Telescope, with an actively controlled segmented single mirror) was discussed by J. Nelson (University of California, Berkeley). There is a site reserved on Mauna Kea, Hawaii, and testing of techniques for producing off-axis parabolic mirror segments is well under way. The testing of these very-long-focal-ratio segments is proving unexpectedly difficult. R. Angel (University of Arizona) described some lessons learned in operating the Multi Mirror Telescope, for example, that low thermal inertia is very important for sharp images, and that separately aimed mirrors are really better than a single completely rigid one, as careful aiming can then partially compensate for large-scale at-

mospheric turbulence. He also showed results of an assortment of techniques for producing thin but very rigid mirrors that might be scaled up to yield either very large or very many reflecting surfaces at reasonable cost. My favourite sandwiched lengths of glass tubing between thin plates, filled the tubes with sand, then gently heated the whole assembly so that the tubes deformed into close-packed hexagonal cylinders of considerably greater strength than round ones.

Among the objects revealed, the star attraction was undoubtedly the 1.57 ms pulsar (see *Nature* 300, 615; 1982), although by the time of the meeting it was already clear to theorists from Berkeley to Bangalore that the very small period derivative means that the object must be old and is not the answer to experimental relativists' quest for a strong coherent astronomical source of gravitational radiation. One intriguing suggestion (J. Arons, University of California, Berkeley) is that the spinning back up to short period occurred by mass transfer from a relatively low-mass companion, which then liberated the pulsar by shedding more than half its mass in a planetary nebula — now seen as the extended thermal source north of the pulsar.

A potpourri of other objects was also discussed: R. Ekers (VLA Observatory) reported that the point infrared source, IRC16, often identified with the centre of our Galaxy, is in fact displaced about 0.1 pc from the compact radio source at the centre of symmetry of a 1-parsec-diameter three-arm spiral source mapped by the VLA. The radio centre is also the centre of symmetry of the highest gas velocities seen near the middle of our Galaxy, and Ekers suggested that IRC16 may be falling into that centre.

W. Hillebrandt (Max Planck Institute, Munich) discussed recent calculations of the collapse of iron cores of massive stars and concluded that core bounce probably ejects material to make a type II supernova and pulsar only in stars of initial masses 8–10 $M_{\odot}$ and not in more massive stars, although nuclear burning in the outer layers, rotation and magnetic fields, which are not included correctly in current models, will all tend to favour ejection and so raise the maximum mass capable of it.

Finally, there were superclusters and the voids between them. A. Szalay (University of Budapest) reminded us that adiabatic fluctuations in the early universe should give rise to very large-scale structures (≥ 150 Mpc and $10^{15} M_{\odot}$), and R. Kirshner (University of Michigan) reminded us that such structures and voids are indeed seen in the distribution of luminous galaxies and clusters, and that, according to Lao-Tzu, "Profit comes from what is there; Usefulness from what is not there". □

Virginia Trimble is Professor of Physics at the University of California, Irvine, California and Visiting Professor of Astronomy, University of Maryland, College Park, Maryland 20742.

*The 10th Texas Symposium, Austin, 13–17 December 1982.

Cell cycle

A major transition for embryos and for embryologists

from Ron Laskey

ONE of the most remarkable features of animal embryos is their speed of cell division. The nuclei of *Drosophila* embryos divide every ten minutes and most of this time is taken up by mitosis rather than DNA replication. Similarly, *Xenopus* embryos replicate their genomes faster than *Escherichia coli* replicates its genome in log phase, even though each *Xenopus* chromosome is about 40 times larger than the entire *E. coli* genome. Of course there are exceptions — such as the sheltered embryos of mammals. By the time the mouse embryo has divided once, the *Xenopus* embryo has already passed 20,000 cells. Recent studies have shown that the periodicity of these short cell cycles is determined by a preset cytoplasmic clock which triggers synchronous cell divisions. In a pair of recent papers, Newport and Kirschner^{1,2} have investigated the transition from this period of rapid synchronous cell divisions in early embryos of *Xenopus laevis* to the longer asynchronous cell cycles characteristic of later embryos or adult cells. They have used an unusual combination of classical embryological techniques with modern biochemical methods to demonstrate and analyse a remarkable sudden transition in many aspects of cell behaviour.

Earlier, using time-lapse cinematography, Hara, Tydeman and Kirschner³ showed that synchronous cleavages in *Xenopus* embryos are accompanied by surface contraction waves and periodic shape changes. Furthermore, the surface contraction cycle persists when cell division is blocked by colchicine, or even when the nucleus is removed.

Newport and Kirschner^{1,2} show that the synchronous cleavage cycles 2–11 occupy 38 ± 2 minutes each. Reports that transcription is switched off throughout this period were confirmed both by gel electrophoresis of labelled RNA and by autoradiography of sectioned labelled embryos. Autoradiography eliminates the possibility that transcription products were not detected simply because an insufficient number of nuclei was analysed. Both criteria revealed that RNA synthesis is switched on suddenly around cleavage cycle 12, the time when the cell cycles elongate and become asynchronous. The onset of transcription was not the only change which correlated with the changing cell cycle. At cleavage cycle 12, cell motility also became detectable for the first time in dissociated cells.

What specifies this sudden transition in cell behaviour? Newport and Kirschner tested several possible models. First, they showed that the timing of the transition does not depend on cytokinesis. Cytochalasin B and gentle centrifugation both block cell division but allow the surface contraction cycles to continue and, more important, they allow the sudden transition to occur at the same time as judged by kinetics of synthesis of DNA and RNA.

An alternative possibility is that the primary event is the switching on of RNA synthesis, and that the changes in cell motility and the cell cycle are direct consequences of new transcription. This explanation was tested by injecting α -amanitin to inhibit transcription by RNA polymerases II and III. The cleavage cycle

and its sudden transition were not inhibited, nor was the sudden onset of cell motility. These events could thus not be consequences of the onset of transcription by RNA polymerases II or III. They are unlikely to be consequences of transcription by RNA polymerase I since transcription products of this polymerase are not detectable at these developmental stages. Instead the transcripts that appear at the transition are mostly tRNA and 5S RNA transcribed by polymerase III, and the small nuclear RNAs which appear to be transcribed by polymerase II. Concentrations of α -amanitin which clearly prevented production of these transcripts did not affect the transition.

If the mid-blastula transition is not a consequence of cytokinesis or new transcription, could it be triggered by measuring elapsed time since fertilization or by counting the number of cycles of nuclear replication? Both these interpretations were ruled out. First, the classical hair ligature experiment of Spemann⁴ was used to delay migration of a nucleus into one half of a partially constricted egg. Cleavage of this half then followed two or three cycles behind the half which originally contained a nucleus. If total elapsed time or number of divisions experienced by a nucleus determines the transition time, then both half-embryos should undergo the transition at the same time. They did not. Instead the half which received a nucleus two cell cycles late underwent the mid-blastula transition two cycles later than the control half. This suggests that the timing of the transition is determined by the ratio of nuclear material to cytoplasm and not by elapsed time or by the number of divisions experienced by a nucleus. A second experiment in which multiple sperm nuclei were introduced reinforced this conclusion.

In the second of their two papers, Newport and Kirschner tested this interpretation directly by microinjecting increasing amounts of pBR322 plasmid DNA into fertilized eggs. If the timing of the transition is determined by the ratio of DNA to cytoplasm, then DNA injection might bring forward the mid-blastula transition. To look at the effect of injected DNA on transcription it was necessary to inject an additional tRNA plasmid to serve as a transcription template, since the amount of endogenous chromosomal template would not be increased by replication. The result was a striking confirmation that the ratio of DNA to cytoplasm determines the timing of the mid-blastula transition. Transcription of the tRNA gene was induced prematurely by injection of pBR322, but only by amounts which equalled or exceeded the DNA content of an embryo at the mid-blastula transition. Presumably the transition is mediated by

100 years ago

BEN NEVIS SUMMIT OBSERVATORY. — Position: 4406 feet above sea, 4m 6f., in the centre of a

rough rocky plateau, covered with felstone lavas and volcanic agglomerates.

From *Nature* 27, 487, 22 March 1883.



Barometer Cairn

Solar Radiation Thermometer

Thermometer Cage

Hut
Standard Rain-gauge

saturation of DNA-binding components of the cytoplasm. These components have not been identified but Newport and Kirschner provide reasonable arguments that they are unlikely to be core histones.

An analogous abrupt switch in the cell cycle has been observed in other amphibian embryos (see refs 1 and 2 for references) and in *Drosophila* embryos, where McKnight and Miller⁵ have shown that it is correlated with the sudden onset of transcription. It will now be interesting to see whether Newport and Kirschner's fin-

dings are widely applicable. Already it is clear that they can have a substantial impact on embryology by integrating several lines of research and formulating clear questions on the relationship between chromosome replication and gene expression in early development. □

1. Newport, J. & Kirschner, M. *Cell* **30**, 675 (1982).
2. Newport, J. & Kirschner, M. *Cell* **30**, 687 (1982).
3. Hara, K., Tydeman, P. & Kirschner, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 462 (1980).
4. Spemann, H. *Embryonic Development and Induction* (Yale University Press, 1938).
5. McKnight, S.L. & Miller, O.L. Jr *Cell* **8**, 305 (1976).

Dynein structure

A delicate molecular bouquet

from Jeremy Hyams

THE beating of the billions of cilia lining your respiratory tract (and, in some cases, oviduct) is driven by the ATPase protein dynein. The dynein molecule forms paired projections, or arms, which extend like rabbit ears from each of the nine doublet microtubules that form the circumference of the ciliary axoneme (the cilium minus its surrounding membrane). By making a transient ATP-sensitive connection with the neighbouring doublet, the dynein arms cause the microtubules to slide over one another, a displacement that is converted into bending by other axonemal structures.

Elucidating the detailed structure of the dynein molecule has long been regarded as essential to understanding its function. The traditional electron microscope (EM) image based on axoneme cross-sections is of a hook¹ which tapers towards its point of attachment to the A-subfibre of the microtubule doublet (see the figure). This is undoubtedly a fairly crude view, however, as a typical EM section is around 80 nm thick and, with arms spaced at intervals of 24 nm, what you see is in fact a superimposition of as many as three different images. Better resolution is obtained from negative staining, although, for reasons that are not entirely clear, this visualizes only the outer of the two arms. These appear as a rod of three subunits² or as a club-shaped projection supported by a thinner spur³.

A radically different interpretation of the structure of the outer dynein arm and its cross-bridge cycle has recently been suggested from replicas of quick-frozen deep-etched axonemes from various ciliated or flagellated cells⁴. The technique gives a spectacular three-dimensional view of the dynein molecule and shows it to be composed of several distinct components. Essentially, these are a pyramid of three globular subunits associated with the A-subfibre and a slender stalk which spans the remaining distance to the B-subfibre of

the next doublet. The relationship between the different elements is reported to show a systematic variation depending on the condition of the axonemes before preparation. In rigor axonemes prepared by the rapid depletion of ATP by dilution and in which the arms are assumed to be physically locked in place⁵, the dynein molecules stand erect, like rows of soldiers. Re-addition of the nucleotide induces the relaxed state in which the arms detach and at the same time undergo a distinct baseward tilt. The basis of this configurational shift is a change in the association between the three globular subunits, and several possible scenarios whereby this could be translated into force-generation have been proposed⁴.

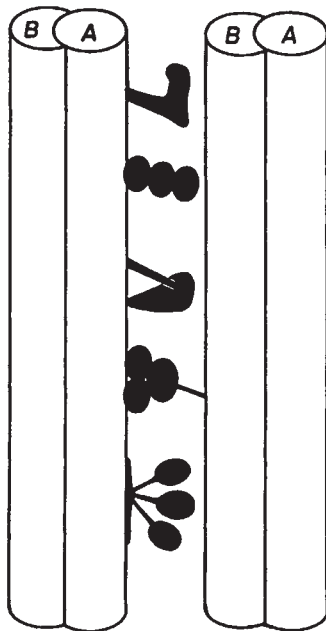


Diagram showing how the EM image of the outer dynein arm can vary with different methods of preparation. From the top: hook (thin section); three-subunit rod (isolated dynein negatively stained); club (*in situ* arm negatively stained); three-head bouquet (deep freeze etch); three-head bouquet (STEM). Each doublet consists of a complete microtubule (the A-subfibre) and an incomplete one (the B-subfibre). (From ref. 6.)

Now, almost before the significance of these novel ideas have had a chance to be fully digested, their validity has been brought sharply into question. In the most recent issue of *The Journal of Cell Biology*, K. Johnson and J. Wall⁶ describe observations, using the scanning transmission electron microscope (STEM) at the Brookhaven National Laboratory, of isolated unstained dynein outer arms obtained from the ciliate *Tetrahymena*, at low electron dose. What they saw was yet another new image of the dynein molecule which they describe as a bouquet of three globular heads connected by separate stalks to a common base. From mass-distribution analysis of dynein molecules that had been rebound to microtubules, Johnson and Wall predict that the root-like base of the dynein molecule forms the permanent structural attachment to the A-subfibre and the globular heads contain the ATP-sensitive sites which interact with adjacent B-subfibres. Johnson and Wall point out that whilst the arrangement they describe is consistent with descriptions of the dynein arm based on other methods of preparation, the same cannot be said of the deep-etched images — although these would be obtained if the stalks of the bouquet collapsed during etching.

Outer dynein arms isolated from the flagella of the green alga *Chlamydomonas* have also been examined under the Brookhaven STEM⁷. Unlike *Tetrahymena* where the outer arm can be isolated as a single particle, the *Chlamydomonas* arm dissociates into two distinct ATPases⁸. The larger of these consists of two heads connected by two strands to a base whilst the smaller appears as a single globular unit. Interestingly, a similar Y-shaped or two-headed dynein arm structure was described some time ago by more conventional EM methods⁹.

Clearly, there is a lot more of the dynein story still to come but it is encouraging that the structure of the ciliary and flagellar arms is at last starting to reflect their biochemical complexity^{10,11}. The fact that several laboratories are currently raising monoclonal antibodies to different arm components should mean that the relationship between structure and function in the ciliary and flagellar axoneme should continue to become clearer. In the meantime, those with a nose for this sort of thing might conclude that Johnson and Wall's bouquet has the right smell about it. □

1. Allen R.D. *J. Cell Biol.* **37**, 825 (1968).
2. Warner F.D., Mitchell D.R. & Perkins C.R. *J. molec. Biol.* **114**, 367 (1977).
3. Witman G.B. & Minervini N. *Symp. Soc. exp. Biol.* **35**, 203 (1982).
4. Goodenough, U.W. & Heuser J.E. *J. Cell Biol.* **95**, 798 (1982).
5. Gibbons, B.H. & Gibbons, I.R. *J. Cell Biol.* **63**, 970 (1974).
6. Johnson, K.A. & Wall, J.S. *J. Cell Biol.* **96**, 669 (1983).
7. Witman, G.B. *et al.* *J. Submicroscop. Cytol.* **15**, 193 (1983).
8. Piperno, G. & Luck, D.J.L. *J. biol. Chem.* **254**, 3084 (1974).
9. Yano, Y. & Miki-Nomura, T. *Biomed. Res.* **2**, 73 (1981).
10. Tang, W.J. *et al.* *J. biol. Chem.* **257**, 508 (1982).
11. Huang, B., Piperno, G. & Luck, D.J.L. *J. biol. Chem.* **254**, 3091 (1979).

Jeremy Hyams is in the Department of Botany and Microbiology, University College London, Gower Street, London WC1E 6BT.

Radioastronomy

Jumps in a pulsar's period

from Andrew Lyne

IN this issue of *Nature*, McCulloch, Hamilton, Royle and Manchester¹ report observations of a large jump or 'glitch' in the period of the Vela pulsar PSR0833-45, followed by a series of smaller changes in period. Over short time periods pulsars usually show remarkably uniform rates of rotation — as indeed would be expected of a spinning body with a large moment of inertia isolated in space. Over longer time periods the rotation rate usually 'spins down' slowly as the pulsar loses angular momentum through the emission of electromagnetic dipole radiation or particles. Irregularities in pulsar rotation, including the jumps observed by McCulloch and his colleagues, are themselves related to the steady spin-down and can tell us a great deal about the pulsar's internal structure.

The most obvious theoretical connection between the changes in period and the slow reduction in rotation rate of a pulsar is through the changes in ellipticity that must occur as centrifugal forces fall. The solid crust of the pulsar must adjust to the changing ellipticity in a series of steps in which the moment of inertia changes and the rotation rate is correspondingly modified abruptly to conserve angular momentum. Such steps might be large enough to explain the glitches observed in pulsars as the Crab pulsar PSR0531 + 21.

For the Vela pulsar, however, the glitches which have been observed so far are so large that they require a change in ellipticity of the order of 1 per cent in a single step. As the ellipticity of the pulsar must be very small to produce its observed rotation period of 0.089 seconds and glitches occur every 2 or 3 years, a different explanation is clearly required. This is provided by a two-component model of the neutron star rather than the single-rigid-body model envisaged above. The outer component is a solid crust while the interior is composed of a liquid which rotates independently, but is loosely coupled to the solid crust. The steady slowdown affects the crust but leaves the interior rotating faster by a fixed amount which depends on the frictional forces coupling the two components. The discontinuity in rotation rate is seen when a sudden increase in coupling between the two speeds up the crust and slows down the liquid interior. Decoupling then follows and differential rotation is slowly re-established. The picture fits very well the pattern of an increase in rotation rate at the glitch and a subsequent exponential decay to steady slowdown over the following few hundred days which has been seen on five

occasions in the Vela pulsar. The changes in coupling between the crust and the fluid interior can be explained by the rather unfamiliar processes caused by vorticity in a superfluid neutron sea and its interactions with the crystal lattice of the crust.

The new element introduced by McCulloch *et al.* is in a series of observations which missed a glitch by only a few hours. As well as the recovery over several hundred days, they show that some of the recovery occurs on a much shorter time scale of only 1.6 days. They interpret this as being due to the effect of a third component, presumably liquid, in the neutron star. The short time scale indicates that it must be closely coupled to the neutron star crust. This may give a new insight into the structure and properties of the exotic material of which neutron stars are composed. Before the nature of the glitches can be understood in detail, however, an observation of the spin-up of

the pulsar during a glitch is necessary.

Clearly the imperfections in periodicity, as observed in the Vela pulsar and many others, are all derived from the steady decrease in the rotation rate². It is not surprising, therefore, that those pulsars which have a small period derivative show a remarkable uniformity of period without glitches or other 'restless' behaviour. In particular, the recently discovered millisecond pulsar PHS1937 + 21 in 4C21.53 with a period of only 1.56 ms has a period derivative of less than about 10^{-19} s⁻¹ so that its period will take more than about 10^9 years to double^{3,4}. This compares with about 10^{15} s⁻¹ for an average pulsar. The resulting stability in rotational period, along with a remarkably well defined and narrow pulse which allows timing accurate to about 1 μ s, may provide us with a clock to surpass the best terrestrial clocks in precision⁴. Unfortunately, unless we find another similar object we will of course never know how good it really is. □

1. McCulloch, P.M., Hamilton, P.A., Hoyle, G.W.H. & Manchester, H.N. *Nature* **302**, 319 (1983).
2. Cordes, J.M. & Helfand, D.J. *Astrophys. J.* **239**, 640 (1980).
3. Ashworth, M., Lyne, A.G. & Smith, F.G. *Nature* **301**, 313 (1983).
4. Backer, D.C., Kulkarni, S.H. & Taylor, J.H. *Nature* **301**, 314 (1983).

Solar flares

The evaporating Sun?

from John C. Brown

IN a recent article in the *Astrophysical Journal* (263, 409; 1982), Acton and his colleagues claim direct evidence of the 'evaporation' of chromospheric matter into the corona of the Sun during a solar flare. They carried out a detailed study of data from the Solar Maximum Mission (SMM) and other well coordinated observations of a solar flare on 7 May 1980. The work constitutes the best evidence yet for evaporation in flares although complications may arise from the sort of transient features in flare energy release reported by Takakura and his colleagues in this issue of *Nature* (see p.317) and the quantitative work cannot yet be considered watertight.

Evaporation is ultimately thought to take place as a consequence of an important property of tenuous astrophysical plasmas — their radiative cooling rates rise steeply up to temperatures of around 10^5 K and then fall at higher temperatures. This occurs largely because of spectrum line emission from trace elements, similar to that causing radiation losses in fusion machines when trace atoms ablate from the walls. The important consequence is that a plasma heated above 10^5 K becomes radiatively unstable, its temperature running away to much higher values until a stabilizing mechanism, such as thermal conduction, sets in to balance the heating. In the Sun, the effect is seen in the separa-

tion of the hot tenuous corona from the cooler denser chromosphere by the thin transition zone which carries the stabilizing conductive flux. If such a plasma is transiently heated, extra mass can become unstable and heat and expand greatly towards new temperature and pressure equilibria. Readjustment of this kind can be important in the explosive behaviour of transients such as supernovae and solar flares.

In the case of the solar flare, the additional heating should depress the transition zone, causing chromospheric matter to heat and expand into the hot coronal regime — a mechanism known widely by the misnomer 'chromospheric evaporation'. It is well established observationally that the transition zone is depressed during solar flares but not until the new work by Acton and his colleagues has there been direct evidence that the chromospheric mass actually 'evaporates' into the corona.

The main data employed by Acton *et al.* comprise hydrogen alpha ($H\alpha$) line profiles from Sacramento Peak Observatory, as indicators of chromospheric structure, and SMM X-ray polychromator (XRP) data on flare coronal structure. SMM hard X-ray burst spectra (HXRBS) and imaging (HXIS) data were used to distinguish possible heating mechanisms driving the evaporation.

By matching the observed $H\alpha$ profiles with profiles computed for a grid of empirical model chromospheres in

Andrew Lyne is in the University of Manchester Nuffield Radio Astronomy Laboratories, Jodrell Bank, Cheshire SK11 9DL.

hydrostatic equilibrium, an estimate was made of the plasma column mass at the top of the chromosphere at points over the flare area. Then, by comparing the flare and preflare values and summing over the flare area, the flare mass leaving the chromosphere was calculated at approximately 10^{14} g. In the coronal regime, XRP data gave a measure of the flare increase in hot plasma emission and this, with the apparent source geometry, allowed estimation of the increase in coronal mass associated with the flare — 5×10^{13} g. The agreement between the estimated chromospheric mass loss and coronal mass gain is within the considerable uncertainties of each, and so is consistent with the evaporation hypothesis. Acton *et al.* further argue that the observed timing and location of broadened H α profiles relative to those of hard X-rays (from HXIS) is consistent with impulsive phase evaporation being driven directly by electron streams; estimates from HXRBS show they have sufficient energy. After the impulsive phase, conductively driven evaporation is assumed to occur.

These conclusions provide *prima facie* quantitative support for the occurrence of chromospheric evaporation in flares. Further support is lent by hard X-ray data indicating increasing coronal mass during flares; observations were made by an occultation technique using two widely separated spacecraft (see *Nature*, **283**, 214; 1980). In this case, though, direct beam heating is not sufficient to drive sufficient evaporation.

There are still areas where the argument for chromospheric evaporation lacks conclusiveness, however. First, the inferred masses are uncertain because of the assumptions made on source geometry and homogeneity. Second, the mass data calculated by Acton *et al.* are integrated over the evaporation period whereas a comparison of rates of mass loss and gain throughout the period would be much more convincing, though very difficult to obtain. Third, the chromospheric models used are in hydrostatic equilibrium whereas, in flare conditions, magnetic pressure in a confining loop is generally important and, moreover, the evaporation process is by its very nature non-hydrostatic. Again, more and more evidence points to the importance of ultra-fast transient features during flare energy release (see Takakura *et al.*, p.317). Finally, and most important, the flare coronal pressure appears greatly to exceed the flare chromospheric pressure even in the post-impulsive stages. This is just contrary to what one would expect for an evaporation process, and suggests that plasma or transient processes are occurring that may invalidate either the model or the radiation diagnostics used. In either case, much further work, both observational and theoretical, will be needed to establish whether and how the chromosphere does evaporate in flares. □

John C. Brown is Reader in Astronomy at the University of Glasgow, Glasgow G12 8QQ.

Atmospheric chemistry

PAN in the natural atmosphere

from Stuart A. Penkett

In this issue of *Nature*, Singh and Salas¹ of the Stanford Research Institute report finding peroxy acetyl nitrate (PAN), an important urban photochemical pollutant, in clean air over the Pacific Ocean (see p.326). Concentrations of PAN varying between 10 and 400 parts in 10^{12} by volume were observed at altitudes up to 8 km. This can only mean that the type of free radical chemistry responsible for photochemical air pollution occurs throughout most of the troposphere².

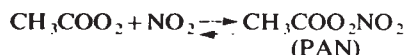
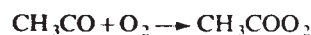
The history of PAN is intimately linked with atmospheric chemistry. It was first discovered in the Los Angeles atmosphere in 1956 (ref.3) and was long considered to be the agent responsible for the unpleasant lachrymatory effects of that city's infamous smog. Twenty years later, the same compound was shown to be the main source of the bleach-like odour perceived when clean linen is hung out to dry in fresh air⁴.

In the intervening years PAN has been found in sunlit urban atmospheres everywhere from Los Angeles to London and from Tokyo to Oslo. The concentration can exceed 10 parts in 10^9 by volume and it is clearly an important photochemical pollutant.

PAN is formed when hydrocarbons are oxidized by hydroxyl radicals in the presence of molecular oxygen and nitrogen dioxide. In the case of propene, a trace gas emitted from the ocean surface, OH addition leads to acetaldehyde formation



then OH abstraction produces the acyl radical and ultimately PAN



Many different types of hydrocarbon, including paraffins, olefines and aromatics give rise to PAN in this way, which makes it a good integrator of chemical activity and, in particular, an excellent indicator of hydroxyl radical chemistry.

A thermal equilibrium exists between $\text{CH}_3\text{COO}_2\text{NO}_2$ and the radicals CH_3COO_2 and NO_2 (ref.5). The position of the equilibrium is such that the molecule is relatively stable at the cooler temperatures experienced above the boundary layer, but is unstable close to the surface. PAN will thus release NO_2 as it is transferred from cooler to warmer temperatures. The NO_2 concentration is only about 30 parts in 10^{12} by volume in the

clean atmosphere⁶ and it has a limited number of sources. The measurements suggest that PAN can act as a reservoir for NO_2 , extending its lifetime and providing a source when others are absent⁷.

Photolysis of NO_2 is the only way of producing ozone in the troposphere, where it may augment the ozone transferred from the stratosphere by general circulation. Production in the troposphere is common in polluted air and is one of the less desirable features of photochemical pollution, but there is so far no clear evidence for ozone formation in clean air. It is worth pointing out, though, that in polluted atmospheres, such as that over southern England, 5 parts in 10^9 of ozone are formed in conjunction with 100 parts in 10^{12} of PAN. This would be equivalent to 10 per cent of the ozone concentration observed in the mid-troposphere.

However it is formed, ozone undergoes tropospheric photolysis with the major consequence that hydroxyl radicals are produced in significant quantities⁸. These radicals can react with many of the different trace gases present in the atmosphere. Trace gases, such as carbon compounds, sulphur compounds, nitrogen compounds and so on, are emitted from a variety of sources, including volcanoes, the biosphere and, more recently, by man's activities. They can have an influence on the atmosphere out of all proportion to their concentration and, in the absence of reaction, would reach levels where effects on climate could occur. Hydroxyl radicals thus have a vital role in maintaining the composition of the atmosphere.

The recent Dahlem Conference on Atmospheric Chemistry⁹ reviewed many of the advances made in this field and drew attention to the notion of a reactive, rather than passive, atmosphere. It is perhaps timely to look in detail at this notion since there is growing evidence that not just carbon dioxide but also several trace gases are increasing in concentration in the atmosphere. This increase may be caused by decreasing removal rates but is more likely associated with increasing emission of materials into the atmosphere. The level of input from natural sources may be changing and pollution is almost certainly a fac-

1. Singh, H.B. & Salas, I.J. *Nature* **301**, 326 (1983).
2. Demerjian, K.L., Kerr, J.A. & Calvert, J.G. in *Advances in Environmental Science and Technology* Vol. 4 (eds Pitts et al.) (Wiley Interscience, 1974).
3. Stephens, F.R., Hansel, P.L., Doerr, R.C. & Scott, W.E. *Ind. Engng Chem.* **48**, 1498 (1956).
4. Lovelock, J.E. & Penkett, S.A. *Nature* **249**, 434 (1974).
5. Cox, R.A. & Roffey, M.J. *Envir. Sci. & Technol.* **11**, 900 (1977).
6. Noxon, J.F. *Geophys. Res. Lett.* **8**, 1223 (1981).
7. Crutzen, P.J. *A. Rev. planet. Sci.* **7**, 443 (1979).
8. Levy, H. *Science* **173**, 141 (1971).
9. *Atmospheric Chemistry*, Dahlem Conf. 1982 (ed. Goldberg, E.D.) (Springer, Berlin).

tor. Large programmes of experiments are now being planned by NASA and by the National Academy of Sciences, in conjunction with the National Science Foundation, with the object of understanding chemical changes in the lower atmosphere and identifying the various processes that impinge

and depend on it. The behaviour of PAN in the natural atmosphere can be expected to be investigated in detail in the near future. □

Stuart A. Penkett is at AERE Harwell, Oxfordshire OX11 0RA.

Solar energy

Towards photoelectrolytic power

from Robert W. Cahn

PHOTOELECTROLYSIS, once an electrochemist's plaything, is close to competing with the photovoltaic silicon cell as a means of converting sunlight into power. This development is due to two kinds of advance: an improvement in the requisite semiconducting photoelectrode and subtle treatment of the electrode surface in order to stabilize it against deterioration. A short while ago, an article in *News and Views*¹ described some of the problems associated with electrodes, especially if they are polycrystalline. One way to solve the problems, described in the same issue², is to use a non-aqueous electrolyte; other possibilities compatible with an aqueous electrolyte, and indeed with water electrolysis, are outlined here.

Many of the key inventions in photoelectrolysis have come from the Bell Laboratories, and so repeat the history of the transistor and the photovoltaic cell. Brattain and Garrett of Bell Labs, in 1955, laid down the now classical theory of the illuminated semiconductor-electrolyte junction. The idea of using an illuminated semiconducting electrode as either anode or cathode in an electrolytic cell to produce electric power or chemical products arose naturally from their work, but the very rapid fall in efficiency of electrodes under irradiation was not solved until 1972, when Honda and Fujishima³ reported, in this journal, the stability of (*n*-type) TiO₂ photoanodes. Unfortunately, such electrodes are useless for converting sunlight as the 3-eV energy gap is much too large. Efficient solar conversion requires gaps in the range 1.3 ± 0.3 eV to absorb light in the λ range 1,000 ± 250 nm.

In 1975, Adam Heller joined Bell Labs to initiate a research programme on photoelectrolysis. At that time it was believed that electrodes capable of giving efficient conversion of sunlight could not do so for long, unless astute chemical counter-measures could be applied at the electrode surface. Success has been achieved at Bell by combining several strategies: photocathodes such as *p*-InP (1.3 eV), *p*-GaAs (1.4 eV) or *p*-CdSe (1.5 eV) are intrinsically stabler than (*n*-type) photoanodes⁴; surface recombination of electrons and holes can be reduced by chemical modification of the electrode surface (anodic formation of a thin hydrated indium oxide layer on InP, or absorption of Ru³⁺ ions at

InP or GaAs surfaces); if electrolysis of water is the aim, the incorporation of thin translucent layers of noble metal helps. In these ways⁵⁻⁷, 'solar' efficiencies of up to 16 per cent have been achieved, well up to photovoltaic levels. Such electrodes can be used either as spontaneous cells (without externally supplied electric power) to make electricity, or in photoassisted externally biased cells to electrolyse water; efficiency can be defined either in terms of electrical power saved or of the power available in an ideal fuel cell from the hydrogen evolved, and the resulting efficiencies differ^{8,9}. One can even combine two photoelectrodes, *p* and *n* type, in a single cell, for example, to electrolyse HBr into H₂ and Br₂ (ref. 9).

Honda and Fujishima's cell used single-crystal TiO₂, and the early Bell Labs cells were also based on single-crystal semiconductors. One of the principal advances at Bell was to establish that the same absorption strategy used to improve electrode kinetics — slow down electron-hole recombination — at a monocrystal surface will also lead to stabilization of grain boundaries against recombination: thus, Ru³⁺ ions improve the efficiency of a polycrystalline *n*-GaAs photoanode by a factor of 4, to the point where such electrodes can be used for practical photoelectrolysis¹⁰. This represents a fundamental advantage over photovoltaics, where it has hitherto proved impractical to slow dopant diffusion along grain boundaries in polycrystalline silicon to the point where good *p*-*n* junctions can be made and the lesser cost of polycrystalline material effectively exploited (see refs 4, 8).

Until recently⁸, it has not been possible to electrolyse water by either spontaneous or photoassisted photoelectrolysis when a photoelectrode with energy gap ~ 1.3 eV has been used. The inefficient *n*-TiO₂ photoanode can, however, be so used¹¹ and is, moreover, highly stable chemically (in spite of being an anode and thus more liable to photocorrosion than a cathode). There is therefore much interest in any procedure which holds promise of increasing the photoelectrochemical efficiency of *n*-TiO₂.

This, it appears, can be done in one of two ways: either by departing increasingly from stoichiometry in the direction of oxygen deficiency — TiO_{2-x} (ref. 12); or by 'alloying' TiO₂ with VO₂ (refs 13, 14). Experiments have been done on both single

crystals and sintered polycrystals, and in these stable materials, grain boundaries cause little deterioration relative to monocrystals (though it is not clear why electron-hole recombination at grain boundaries is insignificant, even in the alloyed materials with lower energy gap). The TiO_{2-x} photoanodes, tested with a K₂SO₄ electrolyte and 335-nm wavelength, increased in efficiency up to 21 per cent for $x = 7 \times 10^{-4}$, then decreased. The decrease was attributed to enhanced surface recombination. It should be pointed out that the recorded 21 per cent efficiency is for monochromatic light, not for a solar spectrum.

The Ti_xV_{1-x}O₂ alloys allow the energy gap to be reduced, down to 1.95–2.04 eV (measurements differ), without, it appears, any loss of chemical stability^{13,14} and without deterioration of efficiency when a polycrystal replaces a single crystal. (Ti_xNb_{1-x}O₂ electrodes were, however, no better than pure TiO₂.) A gap of 2 eV is close to the range for optimum solar efficiency and could be used for a reasonably efficient solar cell. The study of alloyed semiconducting oxide electrodes is only in its infancy, and it may well prove possible to lower the energy gap further without serious loss of chemical stability.

The technological aspects of large-scale photoassisted electrolysis of water have not been examined but it can be most efficiently performed at high temperature, and thus under enhanced pressure, with appropriate electrocatalysts to reduce the hydrogen overvoltage, in devices where multiple layers of anodes and cathodes are interleaved. It is not obvious how such an apparatus could be adapted to permit access of sunlight to the anodes, and if electrolysis is done at ambient temperature and pressure, the loss of efficiency may conceivably outweigh the benefit derived from photoassistance. A study of the combined efficiency of water electrolysis with photoassistance under different conditions would seem to be the next stage in the approach to the capture of sunlight with the aid of semiconductors. □

Robert W. Cahn is Professor of Metallurgy at the University of Paris-Sud, 91405 Orsay Cedex, France.

1. Hamnett, A. & Dennison, S. *Nature* **300**, 687 (1982).
2. Gronet, C.M. & Lewis, N.S. *Nature* **300**, 733 (1982).
3. Honda, K. & Fujishima, A. *Nature* **238**, 37 (1972).
4. Heller, A. *Acc. chem. Res.* **14**, 154 (1981).
5. Heller, A., Miller, B. & Thiel, F.A. *Appl. Phys. Lett.* **38**, 282 (1981).
6. Lewerenz, H.J., Aspnes, D.E., Miller, B., Malm, D.L. & Heller, A. *J. Am. chem. Soc.* **104**, 3325 (1982).
7. Aharon-Shalom, E. & Heller, A. *J. electrochem. Soc.* **130** (in the press).
8. Heller, A. *Solar Energy* **29**, 153 (1982).
9. Levy-Clement, C., Heller, A., Bonner, W.A. & Parkinson, B.A. *J. electrochem. Soc.* **129**, 1701 (1982).
10. Johnston, W.D. Jr, Leamy, H.J., Parkinson, B.A., Heller, A. & Miller, B. *J. electrochem. Soc.* **127**, 90 (1980).
11. Gerstner, M.E. *J. electrochem. Soc.* **126**, 944 (1979).
12. Gantron, J., Lemasson, P. & Maruccio, J.F. *Faraday Disc.* **70**, 81 (1980).
13. Lemasson, P., Etman, M., Gantron, J. & Maruccio, J.F. *Proc. 4th Int. Conf. on Photochemical Conversion and Storage of Solar Energy*, Jerusalem, August 1982 (in the press).
14. Phillips, T.E., Moorjani, K., Murphy, J.C. & Poehler, T.O. *J. electrochem. Soc.* **129**, 1210 (1982).

El Niño Southern Oscillation phenomena

S. G. H. Philander

Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, New Jersey 08540, USA

At intervals that vary from 2 to 10 yr sea-surface temperatures and rainfall are unusually high and the tradewinds are unusually weak over the tropical Pacific Ocean. These Southern Oscillation El Niño events which devastate the ecology of the coastal zones of Ecuador and Peru, which affect the global atmospheric circulation and which can contribute to severe winters over northern America, often develop in a remarkably predictable manner. But the event which began in 1982 has not followed this pattern.

"WHEN pressure is high in the Pacific Ocean, it tends to be low in the Indian Ocean from Africa to Australia." That is how Walker¹⁻⁵ described the large-scale climate fluctuations which he named the Southern Oscillation (SO). Studies during the past half century have confirmed that the SO involves far more than a seesaw in the surface pressure between the South-east Pacific High Pressure Zone and the North Australian-Indonesian Low Pressure Zone (Fig. 1). It is now viewed as one of the most striking examples of interannual climate variability on a global scale: over the tropical Pacific Ocean the SO is associated with considerable fluctuations in the rainfall, the sea-surface temperature, and the intensity of the tradewinds, and it has recently been correlated with droughts over India, with severe winter weather over North America^{6,7}, and with fluctuations of the averaged temperature of the Northern Hemisphere atmosphere⁸.

Early investigators were unable to identify the physical processes responsible for the long-term memory of the atmosphere implied by the persistence, over several seasons, of anomalous conditions associated with the SO. Only in the late 1960s did Bjerknes⁹⁻¹¹ point out that the SO is closely linked with interannual sea-surface temperature variations in the eastern and central Pacific Ocean. He proposed that the ocean, which has a memory much longer than that of the atmosphere, is of central importance in the SO.

Sea-surface temperature anomalies in the tropical Pacific Ocean are primarily associated with a phenomenon known as El Niño. This term originally referred to a warm current that flows southward along the coasts of Ecuador and Peru in January, February and March. The current marks the end of the local fishing season during which sea-surface temperatures are low and the south-east trade winds are intense. (The trades drive the surface waters of the eastern Pacific westward (offshore) and hence induce the upwelling of cold nutrient-rich water near the coast.) In certain years, temperatures are exceptionally high during the warm season and continue to be so during the normally cold upwelling season (Fig. 2a). At present the term El Niño is reserved for those interannual events when anomalously warm surface waters cover not only the coastal zone of South America but also most of the tropical Pacific Ocean. These events have disastrous economic consequences for the fishing and guano industries along the South American coast.

The average period of the SO is 3 yr—spectra of variables such as the surface pressure in the tropical Pacific have a peak at a period of ~38 months—but the SO is so aperiodic that the time between successive El Niño events varies from 2 to 10 yr (ref. 12). It is for this reason that the phases of SO during which sea-surface temperatures are high and the surface-pressure difference across the tropical Pacific is low, are viewed as independent events referred to as El Niño–Southern Oscillation (ENSO) events.

A typical phenomenon

Data sets that describe El Niño–Southern Oscillation events are most reliable for the period since the Second World War.

Of the nine events that have occurred since 1950 seven evolved in a remarkably similar manner. These events—Figs 2 and 3 show their temporal and spatial evolution—started as modest anomalies off the coast of Ecuador and Peru during the boreal spring (February and March) and then expanded westward until the entire tropical Pacific Ocean was affected 6 months later. The minor event of 1963 and the current 1982 event evolved differently because unusually warm surface waters first appeared in the central Pacific Ocean. The coast of South America was affected subsequently so that the phase of the atypical phenomena was very different from the phase of the more common events.

In the following description of a typical event the development is, for convenience, divided into three phases: a precursory phase that precedes the onset off the South American coast in the early spring; next a phase during which anomalous conditions grow; and finally a phase during which anomalous conditions decay and normal conditions return.

(1) Precursors. The westward tradewinds that prevail over most of the tropical Pacific Ocean converge on the North Australian-Indonesian Low Pressure Zone where the air rises and where there is considerable cloudiness and rainfall. The air returns eastwards at greater altitudes and sinks over the cold, dry South-east Pacific High Pressure Zone which is evident in Fig. 1. This zonal cell is known as the Walker Circulation⁹. One of the precursors of El Niño is an eastward displacement of the upward branch of the Walker Circulation, to the region between New Guinea and the dateline. This is revealed by surface pressure, wind and rainfall records¹²: starting in October and November preceding the onset of El Niño the surface pressure at Darwin, Australia, increases, the tradewinds west of the dateline weaken and surface waters near the dateline warm slightly; the rainfall over Indonesia starts to decrease but near the dateline the rainfall increases. This change in the Walker Circulation is also evident in the eastern Pacific Ocean where, at southern latitudes outside the tropics, Easter Island at 27° S–109° W for example, surface pressure starts to fall as early as August preceding the onset of ENSO.

These precursors over the western Pacific appear to be necessary conditions for ENSO events but are not sufficient conditions because similar atmospheric changes can occur even when no El Niño is imminent. During 1979 and 1980 for example, changes in the wind, sea level and rainfall west of the dateline were very similar to those observed during ENSO events but there was no such event over the central and eastern Pacific¹³.

In addition to the zonal Walker cell, the atmosphere has a meridional Hadley circulation. Its salient feature is the Inter-tropical Convergence Zone (ITCZ) where the south-east and north-east trades meet. This narrow zonal band of rising air, cloudiness and high rainfall, which is often visible on photographs taken from satellites, migrates seasonally between 10° N approximately in August and September, and 3° N in February and March. A precursor of ENSO is a southward displacement of the ITCZ which, in the eastern Pacific, is close to or even south of the Equator during the early months of El Niño years.

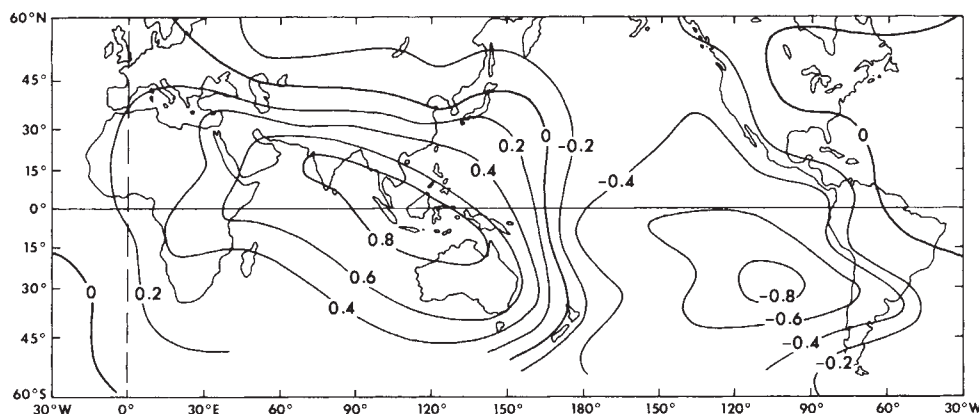


Fig. 1 The correlation of monthly mean surface pressure with that of Djakarta⁴⁵. The correlations are large and negative in the South Pacific High Pressure Zone and are large and positive in the Australian-Indonesian Low Pressure Zone. The SO is not a standing oscillation so that correlations do not have a maximum at zero lag^{46,47}.

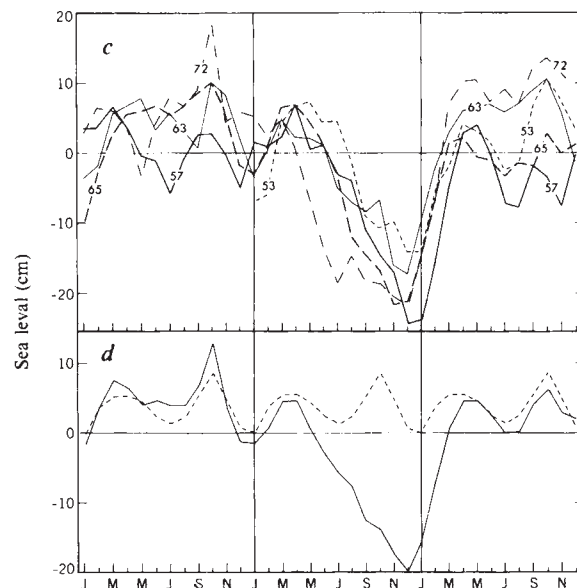
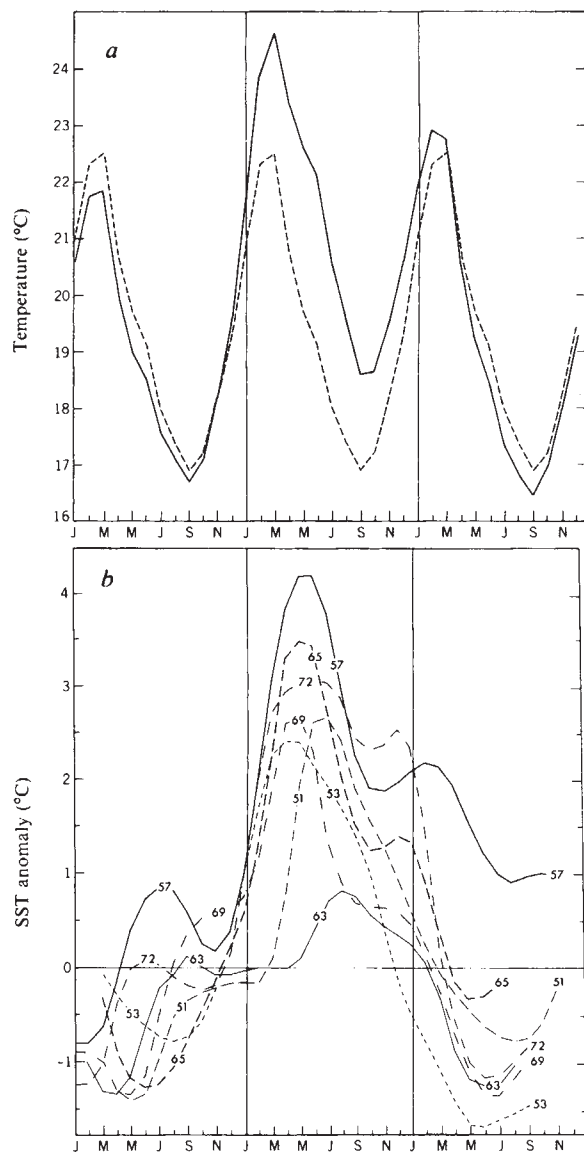


Fig. 2 *a*, Seasonal sea-surface temperature variations (dashed line) and sea-surface temperature variations during a typical ENSO event (obtained by averaging data for the events of 1951, 1953, 1957, 1963, 1965, 1969 and 1972) as measured along a ship track 100 km off the coast of South America between 3° S and 12° S. In the central and eastern tropical Pacific there is a high correlation between sea-surface temperature, sea level and depth of the thermocline, the layer of large vertical density gradients that separates the warm surface waters from the cold deep water of the ocean. *b*, Sea-surface temperature departures from the regular seasonal cycle during the ENSO events listed above, along the ship track that has been mentioned. The major 1976 event, not shown, was similar to the 1972 event. *c*, Sea-level variations at Truk (152° E, 7° N) during the indicated ENSO events⁴⁸. These variations are representative of variations in the western tropical Pacific Ocean where sea-surface temperature changes are small and sea-level fluctuations are highly correlated with fluctuations in the depth of the thermocline. *d*, The average sea-level changes for the ENSO events in *c* (solid line) and the annual variations for years excluding ENSO events (dashed line) at Truk⁴⁸.

This displacement of the ITCZ^{14,15} is associated with weak winds, unusually high sea-surface temperatures and a deep thermocline in the southeastern equatorial Pacific Ocean¹⁶⁻¹⁹. These changes occur early in the calendar year when sea-surface temperatures and the depth of the thermocline in that region are already at a seasonal maximum, when the intensity of the south-east tradewinds has a seasonal minimum, and when the seasonal migrations of the ITCZ take it to its lowest latitude. ENSO during its early stages is therefore an amplification of the seasonal cycle (Fig 2*a*). In the southeastern equatorial

Pacific the changes before and after the onset of ENSO are so similar that it is a matter of definition deciding when the precursors end and ENSO starts.

(2) The growth of anomalous conditions. The most striking feature of the second phase of ENSO is the westward expansion of anomalous conditions that first appear off the coast of Peru and Ecuador (Fig. 3). (The winds within 100 km of the coast of Peru remain normal throughout ENSO but are unrepresentative of conditions further offshore because of the presence of a low level atmospheric coastal jet²⁰.) Because of the westward

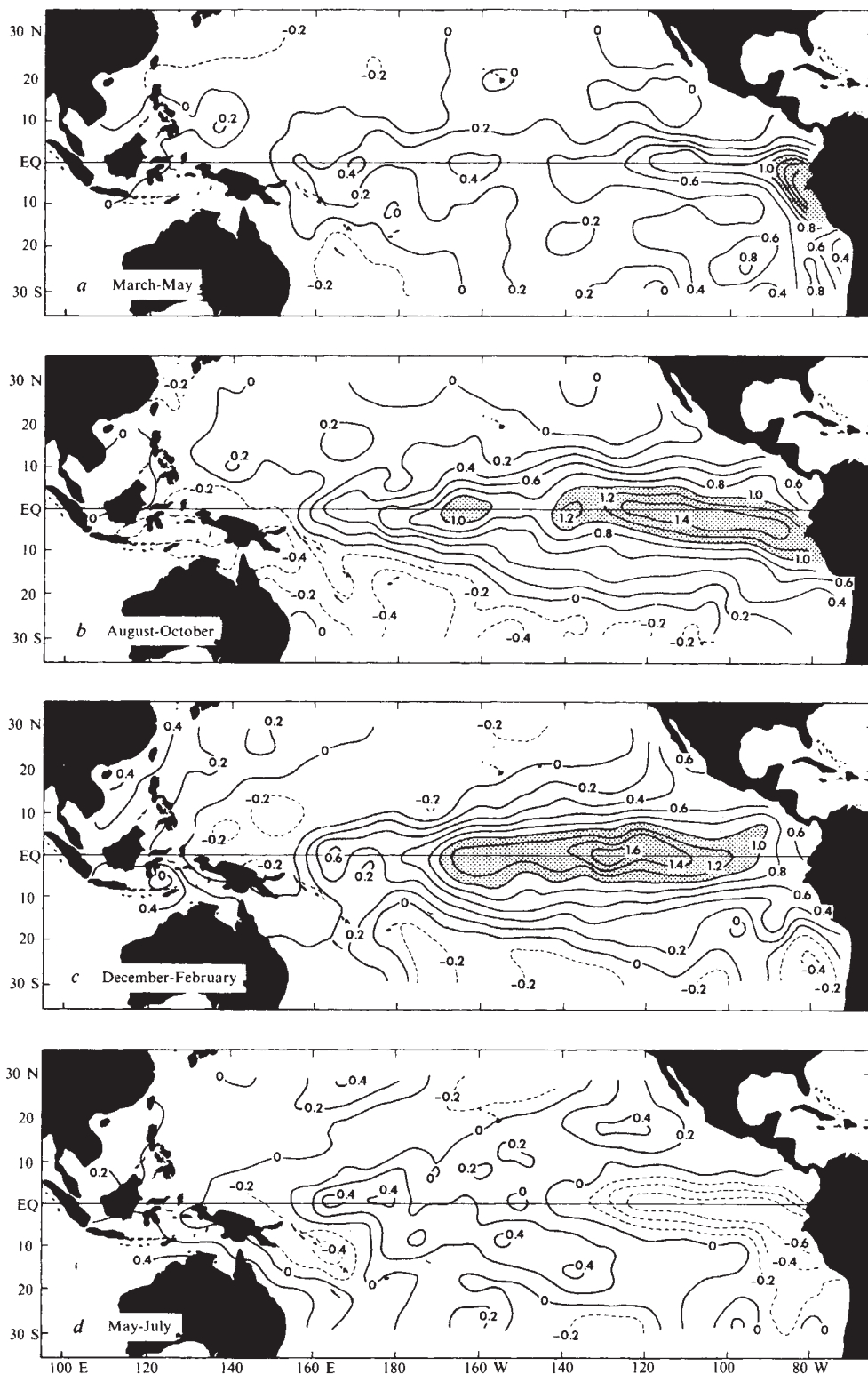


Fig. 3 Sea-surface temperature anomalies (in °C) during a typical ENSO event obtained by averaging data for the events between 1950 and 1973. *a*, March, April and May after the onset; *b*, the following August, September and October; *c*, the following December, January and February; *d*, May, June and July, more than a year after the onset¹².

propagation, at a speed between 50 and 100 cm s⁻¹ exceptionally high rainfall and sea-surface temperatures at Christmas Island (2°N, 157°W) can be predicted several months in advance on the basis of data from the coast of Peru²¹. By October there are anomalous conditions over the entire tropical Pacific Ocean. While the anomalies propagate westwards, the precursors west of the dateline, described earlier, continue to grow in amplitude but their development lags behind the development of anomalous conditions to the east as is evident in Fig. 2 for example.

The Southern Oscillation–El Niño phenomenon is in its mature phase between November and January (after the onset) when there are unusually warm surface waters and exceptionally weak trade winds over most of the tropical Pacific Ocean, when the ITCZ is further south than normal¹⁵ and when the Hadley Cell is intensified²². West of the dateline there are eastward winds which converge on the upward branch of the Walker Circulation which is now in the central Pacific Ocean. Rainfall in the central Pacific is exceptionally high, as is the temperature of the entire tropical troposphere. In the ocean, which is losing

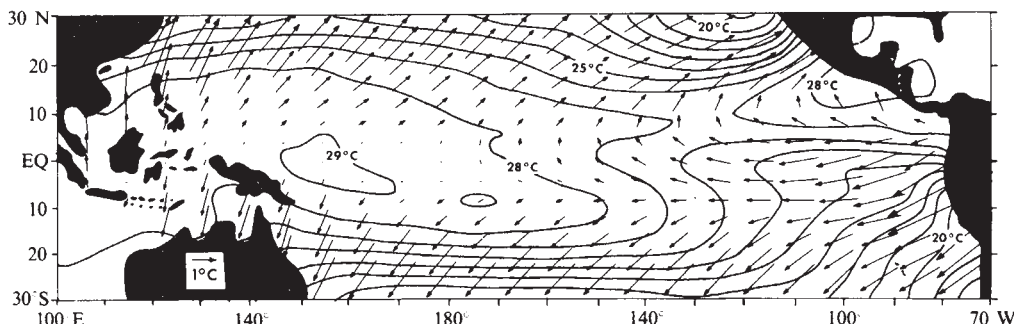


Fig. 4 Mean sea-surface temperatures (solid lines), and the amplitude (length of arrow) and phase (direction of arrow) of the annual cycle of sea-surface temperature in the tropical Pacific Ocean. Maximum sea-surface temperature is in January for an arrow that points downwards and in April for an arrow that points to the left.

heat to the atmosphere, intense eastward currents between the Equator and 10° N carry warm surface waters away from the western Pacific^{16,23} where the depth of the thermocline, and the sea level, decrease sharply (Fig. 2c). Along the western coast of the Americas there is an increase in sea level that propagates poleward in both hemispheres^{24,25}.

In the eastern Pacific, as noted earlier, ENSO is initially an amplification of the seasonal cycle. Even west of 130° W, where the amplitude of seasonal sea-surface temperature variations is very small (Fig. 4), the seasonal cycle continues to modulate the growth of anomalous ENSO conditions. The growth is most rapid when the ITCZ is close to the Equator but is gradual, from July until September, when the ITCZ is furthest from the Equator^{14,15}.

(3) Return to normal conditions. Off the coast of South America the amplitude of anomalous conditions starts to decrease a few months after the onset of ENSO (Fig. 2a, b) but to the west of the Galapagos Islands (0° 90° W) the return to normal conditions does not start until almost 1 yr later. (Along the coast sea level and sea-surface temperature often have a second maximum in January after the onset of ENSO but the amplitude of this signal attenuates rapidly offshore and is small in Fig. 2b, for example). The manner in which anomalous ENSO conditions decay is similar to the manner in which they grow: normal conditions, or sometimes exceptionally low sea-surface temperatures and intense tradewinds first appear in the southeastern tropical Pacific Ocean and then propagate westward until, 12–18 months after the onset of ENSO, normal conditions are restored over the entire tropical Pacific Ocean.

West of the dateline the return to normal conditions follows a different timetable: a decrease in the precipitation at the Line Islands near 160° W, a change in the direction of the winds from eastwards to westwards, and a deepening of the thermocline, start in December and January¹². It is remarkable how, for the different ENSO events in Fig. 2c, the return to normal conditions always starts in the same months.

The atypical 1982 event

In early 1982 anomalous conditions in the western tropical Pacific Ocean were similar to the precursory ENSO conditions described above, namely an eastward displacement of the ascending branch of the Walker cell. During the year the anomalous conditions grew steadily in the western and central Pacific Ocean so that there were, and as of November 1982 still are, severe droughts in Australia and Indonesia, heavy precipitation near the dateline, eastward rather than westward winds over the western Pacific and abnormally high sea-surface temperatures in the central Pacific Ocean. These unusual conditions were confined to the western and central Pacific until the early autumn. Only then did the sea-surface temperature start to increase in the eastern tropical Pacific Ocean. The current El Niño, unlike the more common ones, was therefore not an amplification of the seasonal cycle in the eastern Pacific but was completely out of phase with this cycle.

During November 1982 conditions in the tropical Pacific Ocean were similar to conditions during the mature phase of a typical ENSO event. It will be interesting to see whether or

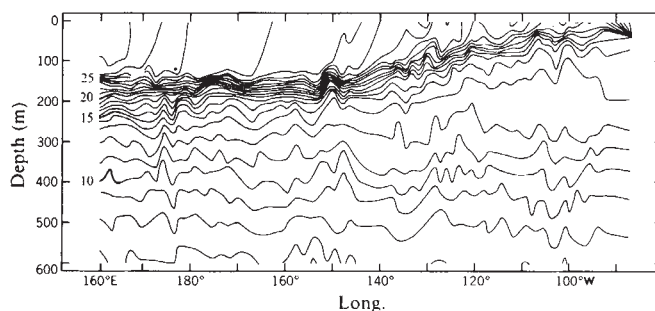


Fig. 5 Temperature (in $^{\circ}$ C) as a function of depth and longitude along the Equator in the Pacific Ocean⁴⁹.

not the further development of the current event resembles that of a typical event.

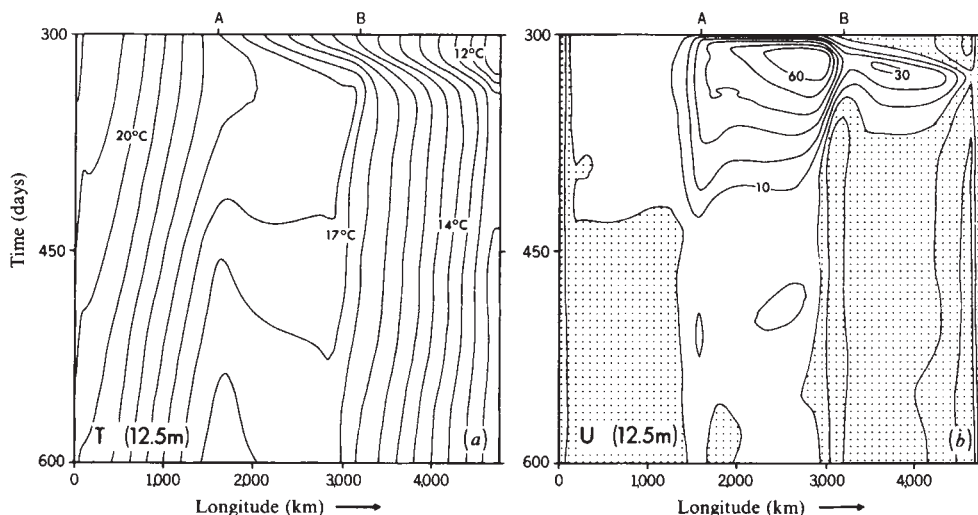
Explanations

Interactions between the ocean and atmosphere are of central importance during ENSO. Anomalous conditions in the ocean and atmosphere develop practically in phase. For an understanding of the phenomenon it is of value to study first two simpler problems: the response of the ocean to the observed meteorological changes; and the response of the atmosphere to the observed sea-surface temperature changes.

(1) Response of the ocean. In the tropical Pacific Ocean the sea-surface temperature has a minimum along the coast of South America and increases in both a westwards and northwards direction (Fig. 4). These temperature gradients exist below the surface too (Fig. 5) and are maintained by the tradewinds. If the winds should stop blowing altogether then a horizontal redistribution of heat will cause horizontal density gradients in the ocean to weaken considerably. In low latitudes this can happen within a matter of weeks or months because the speed of the waves that effect the oceanic adjustment to a change in wind conditions increases with decreasing latitude²⁶. It is for this reason that the intense Somali and equatorial undercurrents in the Indian Ocean, where the monsoons reverse direction, can be generated seasonally, whereas it would take of the order of a decade to generate the Gulf Stream²⁷ (which is at a much higher latitude) from a state of rest.

The weakening of the tradewinds during ENSO will reduce evaporation from the ocean but this cannot explain the anomalously warm sea-surface temperature because the ocean loses an unusually large amount of heat to the atmosphere during ENSO. The exceptionally warm surface waters are apparently caused by a rapid horizontal redistribution of heat in the upper ocean in response to the relaxation of the winds over the Pacific Ocean during ENSO^{16,17}. Several mechanisms, whose relative importance changes during the development of ENSO, are responsible for the redistribution. During the early stages of a typical ENSO the weakening of the meridional component of the tradewinds in the eastern Pacific contributes to the appearance of unusually warm surface waters there. Normally the south-east trades which prevail to the south of the ITCZ cause coastal upwelling and low sea-surface temperatures along the South American coast, but cause downwelling north of the Equator (especially near 3° N) where

Fig. 6 The response of a numerical model of the ocean to spatially uniform westward winds of intensity 0.5 dyn cm^{-2} that suddenly cease to blow between meridians A and B. (The uniform winds had prevailed for 300 days by which time the model ocean was in a state of equilibrium and the most intense current was the sub-surface equatorial undercurrent). The figures show changes along the Equator and near the ocean surface in *a* temperatures and *b* zonal component of the current in cm s^{-1} . (Motion is westward in the shaded region.) Note that the surface waters accelerate eastward temporarily, and become warm permanently, not only between meridians A and B but also in the region to the east³⁵.



sea-surface temperatures are high²⁸. When the south-east trades weaken, warm water flows southwards near the coast. This happens seasonally in February and March²⁹, and also during most El Niño years³⁰. Another factor that contributes to the appearance of anomalously warm surface waters is the weakening of the westward component of the trades. These winds maintain an eastward pressure force, which can be inferred from the zonal temperature gradients in the upper ocean (Fig. 5). When the winds weaken the pressure force is unbalanced and accelerates the warm surface equatorial waters in the west towards the east³¹⁻³⁵. Figure 6 shows a simulation of this effect. The convergent eastward current, which can be observed seasonally in the equatorial Atlantic³⁶ and Pacific³⁴ during the boreal spring when the trades are weak, is augmented during a typical ENSO.

Figure 6 shows that the weakening of westward winds over a confined region will result in a warming not only of that region, but, because of eastward propagating equatorial Kelvin waves, also of the region to the east^{18,33-35}. This warming in the east depends on the zonal extent of the region over which the winds weaken, the magnitude of the change in windstress, and the length of time for which the winds relax³⁵. The amplitude of the changes in the western Pacific in 1982 were apparently sufficient to induce the appearance of exceptionally warm surface waters in the central and, at a later stage, in the eastern Pacific. The striking differences between the 1982 event and the more typical events which are characterized by anomalies that appear off the coast of Peru and then expand westward, suggest that the modest weakening of the trades over the western Pacific during the precursory phase of a typical event is of secondary importance in the development of such an event. A comparison of conditions in the western Pacific in 1979 when a relaxation of the trades failed to induce El Niño in the east, and 1982 when an event occurred, should be fruitful.

Changes in the curl of the windstress also contribute to the redistribution of heat in the tropical Pacific Ocean during ENSO^{16,23}. The curl drives the warm eastward surface current between 3°N and 10°N , and the cold westward current south of 3°N . During most ENSO events the wind changes in such a manner that the eastward current intensifies while the westward current weakens.

Sea-surface temperature changes during ENSO appear to be caused by the weakening of the trades³⁸ with which adiabatic models, forced with the observed winds, simulate thermocline displacements observed during ENSO, confirms this thesis. This, however, is not an explanation for ENSO because the weakening of the trades is part of the atmospheric response to the anomalously high sea-surface temperatures.

(2) The atmospheric response. Exceptionally warm surface waters result in an unusually large release of water vapour to the atmosphere which will be heated should the water condense.

Horizontal temperature gradients in the tropical atmosphere are small so that heating of the lower atmosphere causes rising motion. (In mid- and high-latitudes on the other hand horizontal advection of cold air could balance the heating³⁹.) The rising motion causes the air in the surface layers to converge on the heated region. (The convergence pattern on a rotating globe is complex³⁹⁻⁴³.) Suppose that westward winds are blowing over the heated region. In a downwind direction the air that converges on this region will weaken the westward winds. In an upwind direction the winds will be intensified. In other words, warm surface waters over the eastern tropical Pacific Ocean can cause a weakening of the trades and can even cause eastward winds to appear over the western Pacific. There are, of course, factors other than high sea-surface temperatures that can cause westward winds over the western Pacific.

(3) Air-sea interactions. The arguments advanced above to explain the response of the ocean to anomalous winds, and the response of the atmosphere to unusually high sea-surface temperatures, can be used to determine the fate of a warm sea-surface temperature anomaly that appears in the equatorial Pacific. If the anomaly causes a local heating of the atmosphere then the trades downwind from the anomaly will weaken. Sea-surface temperatures will rise in this region of relaxed winds, for reasons given above. Because of Kelvin waves, warm water will also be transported eastward back towards the original warm anomaly. It follows that this anomaly, in spite of losing latent heat to the atmosphere, will not only persist but will expand in a westward direction. The larger patch of unusually warm surface waters now has an even greater effect on the atmosphere and this in turn causes a further expansion of the warm anomaly. In these unstable air-sea interactions the westward expanding pool of warm water constantly feeds on the warmer water to the west. An anomaly off the coast of Peru, where the sea-surface temperature has a minimum (Fig. 4) will therefore readily expand westwards. It is more difficult for an anomaly in the western or central Pacific to expand eastward because two competing factors affect the sea-surface temperature to the east of an anomaly. One factor is the atmospheric convergence onto the warm anomaly. This causes an intensification of the trades and hence causes enhanced equatorial upwelling and lower sea-surface temperatures. (Westward winds along the Equator induces divergent motion in the upper ocean.) The second factor that influences the sea-surface temperature is the weakening of the winds to the west of the anomaly which can cause a warming of the surface waters to the east as is evident in Fig. 6. Whether or not there is a net increase in the sea-surface temperature to the east of the anomaly depends on the relative importance of the two competing mechanisms. The factors which determine this relative importance, the intensity and shear of the trades for example, require further study.

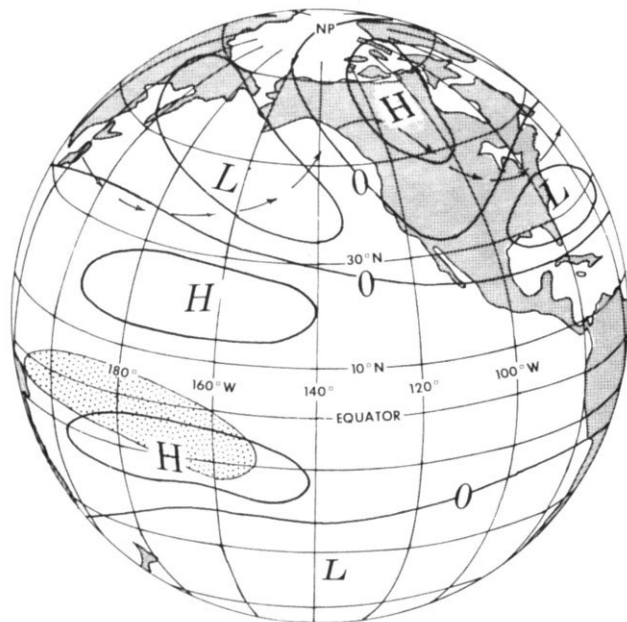


Fig. 7 Teleconnections to the extratropics during ENSO.

In this discussion it has been assumed that unusually warm surface waters heat the atmosphere locally but this is not necessarily so. Suppose that cold dry air descends on a pool of warm surface waters. The pool will lose latent heat to the atmosphere but the heating of the atmosphere will occur non-locally where the air rises and the water vapour condenses. For example, warm surface waters in the southeastern tropical Pacific Ocean where cold dry air descends during most of the year will alter the surface winds in the remote western Pacific where the moist air rises. The warm surface waters, instead of expanding because of the unstable air-sea interactions described above, will therefore disappear.

To grow, a pool of anomalously warm surface waters must affect the atmosphere locally in such a manner that the altered surface winds induce warm oceanic currents towards the pool. This is possible when there is rising air over the pool, a condition which is satisfied in the southeastern equatorial Pacific Ocean when the ITCZ is in its southernmost position, in February and March. This is why the onset of a typical El Niño is always during the boreal spring, and why the subsequent growth of anomalous conditions continues to be modulated by the seasonal cycle. The interplay between the anomalously warm surface waters and the large-scale convergence of the surface winds is very complex because the pool of warm water, especially when it covers a considerable area can modify the large-scale convergence. This happens during the mature phase of ENSO, for example, when the ascending branch of the Walker Cell moves to the central Pacific. In a model of the unstable air-sea interactions described here, the rate at which anomalous conditions grow is sensitive to the relation between the sea-surface temperature and the local heating of the atmosphere. This relation is complex and at present is poorly understood.

The mechanisms that result in the growth of anomalous conditions during a typical ENSO are to some extent self-correcting. During the mature phase, the convergence of surface winds on the ascending branch of the Walker Circulation in the central Pacific is associated with an intensification of the trades over the eastern Pacific where sea-surface temperatures start to fall because of coastal and equatorial upwelling. The attendant reduction in the heating of the atmosphere moves the convergence zone westward so that the restoration of normal conditions gradually progresses in that direction. West of the dateline, as pointed out earlier, the return of normal conditions follows a different course for which there is no theory as yet.

The explanation for the development of ENSO given here has as its starting point the appearance of a modest warm sea-surface temperature anomaly in the eastern tropical Pacific Ocean, or the prolonged appearance of intense eastward winds in the western tropical Pacific. These initial conditions are precursors of ENSO and are associated with an eastward displacement of the upward branch of the Walker Circulation and a southward displacement of the ITCZ. An explanation for these precursors is likely to prove a formidable challenge because the locations of the regions of converging air are determined not only by factors in the tropical Pacific but also by factors elsewhere. Their seasonal migrations, for example, are part of the global seasonal cycle and the precursors of ENSO could be one aspect of an irregular global seasonal cycle. The eastward movement of the ascending branch of the Walker Cell, which was very important in the 1982 event, could be related to changes in the monsoons over the Indian Ocean. As the convergence zones control not only the initiation but also the subsequent development of ENSO a complete understanding of the phenomenon depends on identification of the factors that determine the position of the convergence zones.

Teleconnections with extratropical latitudes

Figure 7 depicts the global teleconnection pattern, at upper tropospheric levels of the atmosphere (heights of 5–10 km) during a Northern Hemisphere winter that coincides with the mature phase of ENSO when unusually warm surface waters cover a large part of the tropical Pacific Ocean. The region of enhanced precipitation in the central Pacific is shaded. The patterns marked H and L are the regions over which the surface of constant pressure at an altitude of ~5 km is anomalously high (H) or low (L). The high pressure over western Canada is associated with unusually low surface temperatures (at Edmonton, Alberta, for example); the low pressure over the southeastern United States is associated with unusually high surface temperatures at Charleston, South Carolina, for example. The arrows indicate how streamlines for airflow in the upper atmosphere are affected.

The correlations between indices of the SO, and certain North American meteorological variables are statistically significant^{6,7,44} only for the Northern Hemisphere winter months of an ENSO year. Furthermore, the correlation coefficients are relatively low so that the tropical Pacific predictors account for less than half the variance of wintertime mean surface temperatures over western Canada for example. In practice this means that not all ENSO events are associated with a severe winter over North America—there was a severe winter during the 1976–77 ENSO event but not during the comparable 1972–73 event—and that exceptionally cold winters can occur even when there are no ENSO events, in 1978 and 1979, for example.

The heating of the tropical troposphere during ENSO events excites large-scale planetary waves in the atmosphere. The ray paths of these waves depend on the mean zonal winds³⁹ and can be such as to cause the perturbations shown in Fig. 7 provided eastward winds prevail from midlatitudes to the equatorial region where the heating occurs. This condition is usually satisfied during the Northern Hemisphere winter months which explains why the observed teleconnections exist in winter only.

Outlook

The available data sets provide a description of a typical El Niño–Southern Oscillation event—this is facilitated by the remarkable degree to which many events are similar—but the data sets are too poor to answer several important questions. Why do some ENSO events attain greater amplitudes than others, and why do some events affect the extratropics but other, comparable events do not? What is the role of anomalous heat fluxes across the ocean surface in the heat budget of the upper ocean? Does the time it takes the ocean to regain

the heat lost to the atmosphere during ENSO determine when the next ENSO event can occur? To what extent is the Indian Ocean involved in ENSO? (Fig. 1 suggests that the Indian Ocean is involved.)

Data gathered over the past few years as part of the US programme EPOCS (Equatorial Pacific Ocean Climate Studies), which is a study of variability on interannual and

shorter time scales in the tropical Pacific Ocean, are beginning to answer some of these questions. This programme will soon be complemented by an international programme that is now being planned.

During this work I benefited from discussions with Drs M. Cane, I. Held, P. Hisard, J. McWilliams, A. Oort and E. Rasmusson who provided preliminary data for the 1982 event.

1. Walker, G. T. *Mem. Ind. met. Dept.* **24**, 75–131 (1923).
2. Walker, G. T. *Mem. Ind. met. Dept.* **24**, 275–332 (1924).
3. Walker, G. T. *Mem. R. met. Soc.* **II**, No. 17, 97–106 (1928).
4. Walker, G. T. & Bliss, E. W. *Mem. R. met. Soc.* No. 4, 53–84 (1932).
5. Walker, G. T. & Bliss, E. W. *Mem. R. met. Soc.* No. 4, 119–139 (1937).
6. Horel, J. D. & Wallace, J. M. *Mon. Weath. Rev.* **109**, 813–829 (1981).
7. van Loon, H. & Madden R. A. *Mon. Weath. Rev.* **109**, 1150–1162 (1981).
8. Pan, Y. H. & Oort, A. H. *Mon. Weath. Rev.* (in the press).
9. Bjerknes, J. *Tellus* **18**, 820–829, 1966.
10. Bjerknes, J. *Mon. Weath. Rev.* **97**, 163–172, 1969.
11. Bjerknes, J. *J. phys. Oceanogr.* **2**, 212–217, 1972.
12. Rasmusson, E. M. & Carpenter, T. H., *Mon. Weath. Rev.* **110**, 354–384 (1982).
13. Donguy, J. R., Henin, C., Morliere, A., Robert, J. P. & Meyers, G. Preprint (Scripps Institution of Oceanography, La Jolla, California).
14. Ramage, C. S. & Hori, A. M., *Mon. Weath. Rev.* **110**, 587–600 (1982).
15. Pazan, E. P. & Meyers, G. *Mon. Weath. Rev.* **110**, 587–600 (1982).
16. Wyrski, K. *J. phys. Oceanogr.* **5**, 572–584 (1975).
17. Wyrski, K. *J. phys. Oceanogr.* **7**, 779–787, 1977.
18. Enfield, D. B. *Resource Management and Environmental Uncertainty* (ed. Glantz, M. H.) (Wiley, New York, 1980).
19. Hickey, B. J. *J. phys. Oceanogr.* **5**, 460–475 (1975).
20. Enfield, D. B. *J. geophys. Res.* **86**, 2005–2016 (1981).
21. Barnett, T. J. *J. phys. Oceanogr.* **11**, 1043–1058 (1981).
22. Reiter, E. A. *Mon. Weath. Rev.* **106**, 324–330 (1978).
23. Wyrski, K. *Science* **180**, 66–68 (1973).
24. Enfield, D. B. & Allen, J. S. *J. phys. Oceanogr.* **10**, 557–578 (1980).
25. Chelton, D. B. & Davis, R. E. *J. phys. Oceanogr.* **12**, 757/784 (1982).
26. Lighthill, M. J. *Phil. Trans. R. Soc.* **A265**, 45–93 (1967).
27. Veronis, G. & Stommel, H. *J. mar. Res.* **15**, 43–69 (1956).
28. Philander, S. G. H. & Pacanowski, R. C. *Tellus* **33**, 201–210 (1981).
29. Sverdrup, H. U., Johnson, M. W. & Fleming, R. H. *The Oceans* (Prentice-Hall, New York, 1942).
30. Wooster, W. S. & Guillen, O. *J. mar. Res.* **32**, 378–404 (1974).
31. McCreary, J. *J. phys. Oceanogr.* **6**, 632–645 (1976).
32. Hurlburt, H. E., Kindle, J. C. & O'Brien, J. J. *J. phys. Oceanogr.* **6**, 621–631 (1976).
33. Cane, M. A. & Sarachik, E. S. *J. mar. Res.* **35**, 2, 395–432 (1977).
34. McCreary, J. thesis, Univ of California, San Diego (1977).
35. Philander, S. G. H. *J. phys. Oceanogr.* **11**, 176–189 (1981).
36. Katz, E. J. & Garzoli, S. *J. mar. Res.* **40**, Suppl., 307–327 (1982).
37. Halpern, D. *J. phys. Oceanogr.* (in the press).
38. Busalacchi, A. J. & O'Brien, J. J. *J. geophys. Res.* **86**, 10901–10907 (1981).
39. Hoskins, B. J. & Karoly, D. *J. atmos. Sci.* **38**, 1179–1196 (1981).
40. Webster, P. J. *Mon. Weath. Rev.* **100**, 518–541 (1972).
41. Webster, P. J. *J. atmos. Sci.* **38**, 554–571 (1981).
42. Gill, A. E. *Q. J. R. met. Soc.* **106**, 447–462 (1980).
43. Keshavamurty, R. N. *J. atmos. Sci.* **39**, 1241–1259 (1982).
44. Barnett, T. P. *Mon. Weath. Rev.* **109**, 1021–1038 (1981).
45. Berlage, H. P. *Kort. Ned. Met. Inst. Meded. Verh.* **69**, 1–152 (1957).
46. Trenberth, K. E. *Q. J. R. met. Soc.* **102**, 639–653 (1976).
47. Chen, W. Y. *Mon. Weath. Rev.* **110**, 800–807 (1982).
48. Meyers, G. *J. phys. Oceanogr.* **12**, 1161–1168 (1982).
49. Colin, C., Henin, C., Hisard, P. & Oudot, C. *Cah. O.R.S.T.O.M. Oceanogr.* **9**, 167–186 (1971).

ARTICLES

On the magnetization and origin of the millisecond pulsar 1937+214

Jonathan Arons

Department of Astronomy, Department of Physics and Space Sciences Laboratory, University of California, Berkeley, California 94720, USA

Lack of an analogue of the Crab Nebula around PSR1937+214 is shown to be consistent with theory of pulsar magnetospheres only if the surface magnetic field is unusually weak, $\sim 10^9$ G. Predictions are made for observations of the weak nebular emission around this object. It is suggested that the neutron star was born as a weakly magnetized, rapidly rotating pulsar, and now is 10^5 – 10^6 yr old.

THE most surprising aspect of the recently discovered¹ millisecond pulsar is the lack of a very bright nebula surrounding the pulsar which emits intense synchrotron radiation in the manner of the Crab Nebula. I show here that modern concepts of the structure of radio pulsars' magnetospheres are simultaneously consistent with the existence of PSR1937+214 as a radio pulsar and with the lack of a readily observable analogue of the Crab Nebula, if the surface magnetic field is unusually weak and if a substantial amount of interstellar extinction exists along the line of sight. The theory constrains the dipole moment of the surface field to lie between $10^{25.7}$ and $10^{26.7}$ G cm³, corresponding to a predicted range of the electromagnetic spin-down rate between 3×10^{-21} ss⁻¹ and 3×10^{-19} ss⁻¹, with a definite preference for the larger values. The main form of rotational energy loss is in a latitudinally inhomogeneous relativistic stellar wind. This excites a surrounding interstellar bubble which emits $\sim 10^{35}$ ergs s⁻¹, mostly in the form of $h\nu \geq 100$ keV X rays. The lack of detection of this bubble in existing observations with the Einstein X-ray observatory is attributed to interstellar extinction, which reduces the flux at Earth below detection threshold in the high sensitivity, low photon energy detectors (especially the Imaging Proportional Counter (IPC)). Predictions are made for further, higher sensitivity observations. I briefly discuss some aspects of this object's origin, with

the conclusion that this star may be an example of a group of short period, weakly magnetized neutron stars which were born in this state and which were not previously recognized because of the lack of sensitivity of the all sky pulsar surveys to objects with periods ≤ 0.1 s.

Constraints on the pulsar's magnetic dipole moment

The most obvious constraint on the pulsar is that the neutron star must be sufficiently well magnetized with a relatively simple external magnetic field, so that it can produce the observed simple radio pulse and interpulse. The magnitude of the required magnetic field cannot be inferred from the radio data without a quantitative theory of pulsar radio emission; no such theory exists. However, the most promising models of the magnetosphere attribute radio emission to be a consequence of relativistic e^+ pair creation above the magnetic polar caps of the neutron star^{2–4}. This happens efficiently if the nondipolar surface components of the external magnetic field are no more complex than low order multipoles, with the strength of the higher multipole moments not greatly exceeding that of the dipole field⁵. Then PSR1937+214 can create a dense pair plasma over its polar caps if its dipole magnetic moment μ exceeds 5×10^{25} G cm³, based on the detailed theory of Arons⁶

and consistent with estimates in earlier work²⁻⁴. This corresponds to a minimum surface dipole field at the poles $\sim 10^8$ G. The total voltage drop across the polar cap is then $\Phi_c \sim 5 \times 10^{14} \mu_{26} P_0^{-2}$ V, with $\mu_{26} = \mu/10^{26}$ G cm³ and $P_0 = P/1.56$ ms. The voltage drop along the field required to create the pairs is much less, $\Delta\Phi_{||} \sim 10^{13} P_0^{-3/8}$ V, independent of μ . Then only about 1 part in 50 of the total electromagnetic energy loss goes into pair creation. The pairs are created by the absorption of ~ 50 GeV γ rays emitted by the curvature process. Lower-energy γ rays do escape from the polar region, but their flux is unobservably small. This model implies $\sim 10^{34} \mu_{26} P_0^{-3/4}$ pairs s⁻¹ are injected into the surrounding environs, in two angular sectors filling about half of the sky as seen from the neutron star with axis of the sectors along the rotation axis of the pulsar. At the pulsar, these pairs have energy per particle ~ 1 GeV, but adiabatic expansion reduces the individual particle energies to nonrelativistic values in the proper frame of the outflow.

The electromagnetic power associated with the wound up magnetic field in this relativistic magnetohydrodynamic wind is $\sim L_{EM}/2$, with⁷

$$L_{EM} \sim \Omega^4 \mu_{\perp}^2 / c^3 \approx 10^{35} \mu_{26}^2 P_0^{-4} \text{ erg s}^{-1} \quad (1)$$

Here $\mu_{\perp}$ is the component of μ perpendicular to the angular velocity Ω . The energy loss in equation (1) corresponds to the minimum spin-down rate

$$\dot{P}_{EM} = 10^{-20} \mu_{26}^2 I_{45}^{-1} P_0^{-1} \text{ s s}^{-1} \quad (2)$$

Recent theories⁸⁻¹⁰ suggest the electromagnetic torque vanishes if the magnetic moment is aligned, so that equation (1) is the same as the energy loss for the vacuum rotator^{11,12}. A more realistic theory of the magnetosphere¹³, in which the region around the magnetic equator has electromagnetic energy loss similar to that of a large amplitude wave with some characteristics of the vacuum theory while the rotational polar regions lose energy in the magneto hydrodynamic wind, suggests a different scaling of L_{EM} with Ω than is given in equation (1), but with almost the same results for $\mu_{\perp}$, when the period is short and the light cylinder distance $R_L = c/\Omega$ (here $75 P_0$ km) is not too much greater than the stellar radius $R_* \sim 10$ km. $R_L/R_* < 10^2$ is sufficient to allow the use of equation (1) as a reasonable approximation.

The equatorial sector is much more effective in accelerating particles. In this region, ions are drawn up from the surface to balance the electron current flowing out through the pair producing regions. They have a density $n_i(r \sim R_L) \sim \Omega B(R_L)/2\pi c e \sim \Omega^4 \mu/2\pi c^4 e$. Electrons and positrons with similar density are drawn into this zone from the boundaries of the polar wind sectors. The electromagnetic wave field cannot coexist stably with plasma of this density¹⁴⁻¹⁶. Present evidence¹⁷ suggests the energy is converted fairly completely into particle acceleration. The ions are accelerated to the energy $\sim e\Phi_c \sim e\mu\Omega^2/c^2 \sim 5 \times 10^{14} \mu_{26} P_0^{-2}$ eV. They contribute $L_i \sim L_{EM}/4$ to the total power injected into the outside world. This energy per ion is $\sim 10^2$ greater than would be estimated from the theory of particle acceleration in strictly vacuum waves¹², as the field here has a phase velocity well above c and therefore approximate phase locking does not occur. The electrons and positrons are similarly accelerated, but curvature radiation reaction limits their energy to $\gamma_p mc^2 \sim (\Omega\mu/c^2)^{1/4} mc^2 \sim 7 \mu_{26}^{1/4} P_0^{1/4} \times 10^{12}$ eV. This region of the magnetosphere then emits a broad spectrum of γ rays extending up to $\sim 2 \mu_{26}^{3/4} P_0^{1/4}$ GeV, modulated at the radio pulse rate with a luminosity $\sim L_{EM}/4$ also. This implies a pulsed γ -ray flux at Earth above 100 MeV $\sim 3 \times 10^{-8} \mu_{26} P_0^{-4} D_2^{-2}$ photons cm⁻² s⁻¹, where D_2 is the distance measured in units of 2 kpc. The beaming factors are unknown, but this certainly suggests a careful search among the fainter COS B sources might be fruitful, if μ_{26} exceeds 3.

X-ray emission from the surrounding region provides more stringent constraints on μ . The relativistic wind in the equatorial sector is dominated by the kinetic energy of the ions, with a

total number loss $\sim 10^{32} \mu_{26} P_0^{-2} \text{ s}^{-1}$. This wind drives a bubble into the surrounding medium, with characteristics similar to those described by Rees and Gunn¹⁸ in their model of the Crab Nebula.

(1) Supersonic expansion: If the pulsar is sufficiently young, the bubble expands into the surrounding medium, driving a shock wave with radius $R_s \sim (L_{EM}/8\pi\rho)^{1/5} t^{3/5}$, where ρ is the mass density of the external medium and t the time since the formation of the pulsar. A density $\rho \sim 10^{-27} \text{ g cm}^{-3}$ might typify either the ambient interstellar density or the density of the ejecta well inside a young supernova remnant. Then $R_s \sim 0.7 \mu_{26}^{2/5} n_3^{1/5} P_0^{-4/5} t_3^{3/5}$ pc, with $t_3 = t/10^3$ yr. The pressure in the shocked material is $p_s \approx 3 \times 10^{-12} n_3^{3/5} \mu_{26}^{4/5} P_0^{-8/5} t_3^{-4/5} \text{ dyne cm}^{-2}$. If $t_3 \gg 1$, this model predicts the pressure to fall below the ambient interstellar pressure $\sim 10^{-12} \text{ dyne cm}^{-2}$, the bubble then expands subsonically, a case treated below and the circumstance likely to be of most interest to the observed pulsar.

The shocked external medium lies in the annulus $0.85 R_s < r < R_s$, with an utterly unmeasurable amount of bremsstrahlung X-ray emission. For $r < R_w = (L_{EM}/8\pi\rho c)^{1/2}$, the unimpeded relativistic wind flows out from the pulsar. At $R_w \sim 0.07 \mu_{26}^{3/5} P_0^{-6/5} n_3^{-3/10} t_3^{2/5}$ pc, a relativistic shock wave can decelerate the flow, reheating the electrons and positrons to random energy per particle comparable with that of the ions^{18,19}. Then a nonthermal synchrotron nebula is formed, with synchrotron emission occurring for $R_w < r < R_b$ and with total X-ray luminosity $\sim L_{EM}/4 \approx 10^{34} \mu_{26}^2 P_0^{-4} \text{ erg s}^{-1}$. The spectrum depends on details of the particle distribution accelerated at the shock. If it is sufficiently flat, the emitted spectrum is dominated by the highest energy particles, as if the pairs were all accelerated to the single energy $10^9 \mu_{26} P_0^{-2} m_e c^2$. The magnetic field just downstream from the shock is $B(R_w) \sim 2 \times 10^{-5} n_3^{3/10} \mu_{26}^{2/5} t_3^{-2/5} P_0^{-1/5} \text{ G}$ and increases linearly with radius, while the flow velocity decreases from $c/3$ at the shock, with the decrease in proportion to r^{-2} . Then the spectral luminosity is $L(\epsilon) = (L_{EM}/4\epsilon_1)(\epsilon/\epsilon_1)^{-1/2}$ in the range $\epsilon_2 < \epsilon < \epsilon_1$, with $\epsilon_1 \sim 400 n_3^{3/2} \mu_{26}^{12/5} t_3^{-2/5} P_0^{-24/5} \text{ keV}$ and $\epsilon_2 \sim 0.01 n_3^{-6/5} \mu_{26}^{-3/5} P_0^{6/5} t_3^{-7/5} \text{ keV}$. For $\epsilon < \epsilon_2$, the spectrum decreases proportional to $\epsilon^{-1/3}$. This translates into a 2–10 keV flux at Earth $\sim 5 \times 10^{-12} \mu_{26}^{4/5} n_3^{3/4} P_0^{-8/5} t_3^{1/5} D_2^{-2} \text{ erg cm}^{-2} \text{ s}^{-1}$. If $\mu_{26} > 3 D_2^{5/2}$, this should have been detected by the monitor proportional counter on the Einstein satellite during existing observations of this region (J. E. Grindlay, personal communication). A flux of this magnitude for all acceptable values of $\mu D^{-5/2}$ was easily observable in the IPC, if the neutral column density was $< 10^{22} \text{ cm}^{-2}$. However, the distance indicated by the dispersion measure $D > 2$ kpc puts the pulsar in a region where stars with a distance > 1.5 kpc show 2–4 mag of extinction (D. van Buren, personal communication), corresponding to a neutral hydrogen column in excess of 10^{22} cm^{-2} and an absorption cutoff such that this source would have been invisible in the IPC. This amount of obscuration is consistent with the red colour observed in the optical candidate²⁰ even if the input spectrum were as flat as $f_\nu \propto \nu^{1/3}$. The nebular optical flux predicted by this model corresponds to an unobserved optical magnitude $m_v > 21$, which is unobservable if interstellar extinction exceeds 2 mag by any except heroic efforts.

(2) Subsonic expansion: If the system is old ($t_3 > 1$), the bubble expands subsonically with internal pressure comparable to the ambient pressure. This pressure is most likely to be comparable to the normal interstellar pressure $p_0 \sim 10^{-12} p_{12} \text{ dyn cm}^{-2}$. Then the outer radius of the bubble is $R_b \sim (3L_{EM}/64\pi p_0)^{1/3} t^{1/3} \approx 2.5 \mu_{26}^{2/3} p_{12}^{-2/3} t_3^{1/3} \text{ pc}$, where $t_4 = t/10^4$ yr. The relativistic wind in the equatorial sector is thermalized at $R_w = 0.14 \mu_{26} p_{12}^{-1/2} \text{ pc}$. Synchrotron X rays come from the inner region of the bubble $r \sim R_w$. The same assumptions as were made in the supersonic case lead to an X-ray flux proportional to $\epsilon^{-1/2}$, with upper cutoff at $\epsilon'_1 \sim 170 p_{12}^{1/2} \mu_{26} P_0^{-4} \text{ keV}$ and a 2–10 keV flux at Earth $\sim 10^{-11} \mu_{26} p_{12}^{-1/4} \text{ erg cm}^{-2} \text{ s}^{-1}$. The limits on the optical emission are similar to the supersonic case.

If the pulsar moves supersonically with respect to the interstellar medium and its age exceeds $\sim 10^3$ yr, static pressure

confinement is replaced by ram pressure confinement. In the direction of the pulsar's motion, the relativistic wind is still terminated by a shock wave at the radius R_w given above, but with $p_{12} = (\rho v^2 / 10^{-12} \text{ dyn cm}^{-2}) = 100(n/1 \text{ cm}^{-3})(v/100 \text{ km s}^{-1})^2$. Here v is the speed of the star with respect to the medium and ρ is the mass density of the interstellar medium. The spectral characteristics of the shock heated electrons and positrons are the same as before, but the increased value of p_{12} moves the upper cutoff of the spectrum to higher energy and therefore reduces the X-ray flux in a fixed band in proportion to $p_{12}^{-1/2}$. The bubble is now brightest at the apex of the pulsar's motion, with an elongated 'contrail' structure downstream from the star.

Radio synchrotron emission might come from the pairs injected in the magnetically dominated wind originating from the rotational polar regions of the magnetosphere. The amount of reacceleration of the pairs is unclear, because the wind in this region is magnetically dominated, and the jump conditions for a shock at R_w indicate very little particle acceleration at the shock. This wind must decelerate more gradually than the plasma dominated flow in the equatorial zones. In the absence of a physical theory of particle acceleration, I argue through analogy with the Crab nebula, where the total energy in relativistic electrons (and positrons, most likely) is comparable to the magnetic energy contained and whose X-ray²¹ and radio emission^{22,23} in the inner regions has the sector structure consistent with the injection pattern of particles and fields outlined here. If similar equipartition is assumed for the bubble blown by PSR1937+214, and a similar power law spectrum of relativistic electrons and positrons is assumed, with the number of particles with energy between E and $E + dE \propto E^{-3/2}$, then the radio flux expected from the bubble is

$$f_\nu \approx N c \sigma_T p_0 / 3 \nu_c^{3/4} \nu^{1/4} (4\pi D^2) \\ \approx 200 \mu_{26} p_{12}^{5/8} P_0^{-3/4} t_5 \nu_9^{-1/4} D_2^{-2} \mu\text{Jy}$$

where ν_9 is the frequency/1Ghz, N is the total number of electrons and positrons injected since the pulsar began producing the relativistic outflow $\approx 3 \times 10^{45} \mu_{26} P_0^{-3/4} t_4$, σ_T is the Thomson cross-section, $\nu_c = 12.6 p_{12}^{1/2} \text{ Hz}$ is the cyclotron frequency and t_5 is the time since the pulsar began emitting the relativistic wind per 10^5 yr . I argue below that $t \sim 10^5 - 10^6 \text{ yr}$. This flux is undetectable if $t_5 D_2^{-2} < 10^2$, as is the IR and optical flux from this part of the wind. The application of the same calculation to the Crab nebula yields a radio flux $\sim 800 \nu_9^{-1/4} \text{ Jy}$ at a distance of 2 kpc, close to the observed 1 kJy. The difference between the results occurs because: (1) the magnetic moment of the Crab pulsar is $\sim 10^4$ times larger; (2) the confining pressure is $\sim 10^4$ times greater (this may reflect confinement of the Crab bubble by a shock driven into the surrounding medium¹⁹); and (3) absorption of synchrotron γ rays from newly created pairs in the magnetic field just above the neutron star's polar cap enhances the number of pairs created per unit time by a factor ~ 50 above what would be predicted by the estimate used here, which is based on absorption of curvature photons only. This last difference occurs because of the much greater opacity for γ -ray conversion in the stronger magnetic field of the Crab pulsar.

Thus, in spite of the rapid rotation rate observed for this pulsar, the absence of a detectable, nonthermally emitting nebula around the star is still consistent with current ideas about the electrodynamics behind the pulsed radio emission, if the magnetic moment is $> 0.5 \times 10^{26} \text{ G cm}^3$ but does not exceed $\sim 5 \times 10^{26} \text{ G cm}^3$, corresponding to a dipolar surface field in the range $10^8 \text{ G} < B_{\text{pole}} < 10^9 \text{ G}$. This gives a minimum spin-down rate in the range $3 \times 10^{-21} I_{45}^{-1} < \dot{P} < 3 \times 10^{-19} I_{45}^{-1} \text{ s s}^{-1}$. Since the pulsar has a well defined pulse structure and has not shown any sign of nulls²⁴, a dense magnetosphere is indicated, implying the pulsar is not close to the pair creation threshold. Therefore, the theory suggests $\dot{P}$ is not close to its observationally acceptable lower bound, $\dot{P} \geq 10^{-20}$. After this paper was submitted, $\dot{P}$ was measured²⁴ to be $\approx 10^{-19} \text{ s s}^{-1}$, implying $\mu \approx 3 \times 10^{26} I_{45}^{1/2} \text{ G cm}^3$, within the range of acceptable electromagnetic torques found

here. A positive test of the relativistic wind theory would come from the detection of a nebular source with properties like those outlined above.

Possible origins for PSR1937+214

The facts which bear on the origin of the millisecond pulsar are (1) it lies close to the galactic plane, $b = -0.3^\circ$; (2) it has a small spin period $P = 1.56 \text{ ms}$, close to the maximum possible for a rotating neutron star; (3) its dipole magnetic field is unusually weak in comparison to the previously known neutron stars; (4) no shell-like supernova remnant is observed near the object¹. The spin-down time $P/2\dot{P} \sim 2 \times 10^8 \text{ yr}$ is not a reliable indicator of the age, since the initial spin-down time $t_i = 10^9 (P/1 \text{ ms})^2 \mu_{26}^{-2} I_{45} \text{ yr}$ is comparable with or greater than $P/2\dot{P}$ for acceptable values of $I_{45} \mu_{26}^{-2}$.

One plausible possibility is that the neutron star was born as a radio pulsar, presumably in a supernova explosion. The absence of an obvious supernova remnant driven by the blast wave of the explosion requires the pulsar to be $> 10^5 - 10^6 \text{ yr}$ old. If the component of the star's space velocity normal to the galactic plane is $\sim 10^2 \text{ km s}^{-1}$, as is typical of the whole pulsar population²⁵, $b = -0.3^\circ$ now implies the age t and the transverse velocity $v_\perp$ satisfy $(v_\perp/100 \text{ km s}^{-1})(t/10^6 \text{ yr})(5 \text{ kpc}/D) < 1$, where I have assumed the pulsar was born $\sim 50 \text{ pc}$ on the opposite side of the galactic plane in order to find an upper bound for t . Recent 21-cm studies suggest the pulsar's distance is $\sim 5 \text{ kpc}$ (D. C. Backer and S. Kulkarni, personal communications). Thus the absence of a supernova remnant is barely consistent with the star having been born as an isolated object. It must have been born with an unusually small magnetic moment. A large magnetic moment implies a spin-down time much less than the currently observed $P/2\dot{P}$ and therefore an initial period much less than the present period, which is impossible. Therefore $t_i \geq t \sim 10^6 \text{ yr}$ (assuming $v_\perp \sim 100 \text{ km s}^{-1}$), which yields an initial $\mu \leq 10^{28} (10^5 \text{ yr}/t)^{1/2} I_{45}^{1/2} \text{ G cm}^3$. This limit can be relaxed if the pulsar's space velocity is much smaller than usual but the increase is unlikely to be greater than a factor of 3, since $v_\perp$ is unlikely to be less than the one-dimensional velocity dispersion of massive stars $\sim 10 \text{ km s}^{-1}$.

Such small magnetic moments are unusual, as the existing pulsars probably were born with $\mu \sim 10^{30} \text{ G cm}^3$ and with relatively long rotation periods²⁶. All pulsar models which rely on electron-positron pair creation imply that objects with relatively long rotation period and weak field ($\mu < 10^{29} \text{ G cm}^3$) would never appear as pulsars, while the all sky surveys are very insensitive to the detection of objects with periods $< 0.1 \text{ s}$ (ref. 27 and J. H. Taylor, personal communication). Weak field cores of pre-supernova stars may be weakly coupled to the stellar envelope and thus might preserve the angular momentum per gramme typical of the upper main sequence even during the evolution of the stars into the giant and supergiant phase. Thus a weakly magnetized core could give rise to a weakly magnetized, rapidly spinning neutron star (initial $P \sim 1 \text{ ms}$), while the more strongly magnetized presupernova cores give rise to more slowly rotating pulsars, since their angular momentum is magnetically transferred to the expanding envelope. The weakly magnetized neutron stars then can appear as pulsars. They obtain a pulsar like polar cap voltage drop ($\Phi_c > 10^{14} \text{ V}$) because their rapid rotation makes up for their weaker magnetization. The pulsar surveys therefore are likely to have fully sampled the long period pulsars, but there may be an interesting number of short period objects in the galactic plane waiting to be discovered which would fill in the presently underpopulated $P < 0.1 \text{ s}$, $\dot{P} < 10^{-15} \text{ region of the } P-\dot{P} \text{ diagram}$. If this is the origin of PSR1937+214, it suggests magnetic coupling of the presupernova core to the envelope breaks down when the magnetic flux threading the core is $\leq 10^{22} \text{ Mx}$ (surface field of the core less than $\sim 10^4 \text{ G}$), assuming the magnetic flux is frozen in during the core collapse to the neutron star.

Rotational spin up during accretion²⁸ is an alternative origin for the short period, if the pulsar was once in a binary system. In this interpretation, the neutron star and its magnetic field

are older than the radio pulsar. This opens the possibility that the neutron star had 'normal' magnetization ($\mu \sim 10^{30} \text{ G cm}^3$) at its birth, but since has undergone decay of its dipole moment (and of all large scale components of the surface field) to the present low value. To achieve a rotation period as short as $\sim 1 \text{ ms}$, the magnetosphere must be compressed by the infalling matter until the magnetopause radius during accretion is less than the keplerian rotation radius corresponding to the presently observed period, a constraint noted independently by Alpar *et al.*²⁹ in an article I saw after the present paper was submitted. The theory of accretion onto magnetized neutron stars³⁰⁻³² then requires an accretion luminosity $L_a > 4 \times 10^{38} (\mu/3 \times 10^{26} \text{ G cm}^3)^2 (M/1 M_\odot)^{-7/6} \text{ erg s}^{-1}$ during the mass transfer phase, with a total mass transferred $\sim 0.1 M_\odot$ required to reach the present period. The basic problem with this sort of scenario is the long time needed to transfer the mass. If L_a is not greater than that of the brightest sources in the Magellanic clouds, $\sim 10^{39} \text{ erg s}^{-1}$, the time needed for mass transfer is $\sim 10^7 \text{ yr}$, greatly in excess of any rapid mass transfer phase in high mass, close binary systems with conservative mass transfer³³. More rapid mass transfer rates may be possible, as in the peculiar system SS433, but their duration is likely to be still shorter than in the normal X-ray pulsars. If the binary is a low mass system with the neutron star as the more massive companion, as in the scenario of Alpar *et al.*²⁹, the stars recede from each other during conservative mass transfer in spite of the effects of gravitational radiation³⁴. Times as long as 10^7 yr for mass transfer with $\sim 0.1 M_\odot$ transferred are now possible, but the final state is a neutron star still bound to the low mass companion with an orbital period of a few days. The timing data completely exclude such a bound state for PSR1937+214. If the mass transfer is strongly nonconservative, the stars might approach each other and merge, which would lead to rapid final spin. However, in low-mass systems, this fate is unlikely, as mass transfer begins in the conservative mode and causes the stars to recede, not merge, with the whole process remaining in the conservative mode. Mass transfer in a wide binary (cases B and C of Kippenhahn and Weigert³⁵) has similar defects. In particular, even a wide, low-mass binary with separation $\sim 1 \text{ AU}$ (which is close enough to have had a rapid mass transfer phase in the past) would have $\dot{P}_{\text{orbit}} \sim 10^{-16}$, and this is already excluded by the several days of timing which yield $\dot{P} \sim 10^{-19}$. Therefore I conclude that no simple mass transfer model with a sufficiently large amount of angular momentum and mass transfer exists. Even if an acceptable model is found, the location of the star close to the galactic plane constrains the age of the pulsar in the same manner as is true for a lonely birth. In addition, any mass transfer model does not easily explain the small value of the galactic latitude, since the long spin-up time implies a greater height above the plane even if the binary were massive and began its life in the plane. In the low mass binary idea, with the system drawn from the spherical bulge population, the location near the plane is a complete accident. Thus, I prefer the picture of an isolated birth with weak initial magnetization and rapid initial rotation, implying weak coupling of the weakly magnetized presupernova core to the envelope of a high mass star. Then the pulsar's age probably is $< 10^6 \text{ yr}$, and a substantial number of such objects, with $P \ll 0.1 \text{ s}$ and $\dot{P} < 10^{-15} \text{ s s}^{-1}$, may exist near the galactic plane.

Conclusions

The most striking feature of PSR1937+214 is the lack of an analogue of the Crab Nebula around it. I have argued that this star is an example of a normal pulsar, unusual in its short rotation period and magnetization but normal in the polar cap voltage available and therefore in its gross radio emission properties. For any $\dot{P}$ smaller than 3×10^{-19} , an electromagnetic spin-down interpretation is tenable, but if $\dot{P}$ were smaller than 3×10^{-21} , a substantial conflict would occur with modern ideas about the dense electron-positron plasma likely to be in the magnetosphere and the relativistic wind which carries the angular momentum away from the star. Thus it is comforting that observations²⁴ made after submission of the present paper indicate $\dot{P} \sim 10^{-19}$ consistent with the electromagnetic interpretation given here and by Alpar *et al.*²⁹. A search for the X-ray emission from the wind driven bubble is likely to yield positive results at photon energies above a few keV and at the level of $\sim 0.01\text{--}0.1$ Uhuru counts, if the distance is $\sim 5 \text{ kpc}$. The star may have been born with weak magnetization and rapid initial spin. If this is so, the (difficult) search for short period pulsars ($P \ll 0.1 \text{ s}$) would yield positive results mostly in the galactic plane.

No matter what the interpretation of the star's origin, the small magnetization inferred for this star is quite interesting. Alignment is not a likely explanation of the low observed electromagnetic torque, as the pulsar exhibits a well defined pulse and interpulse. While this structure may be understandable in terms of beaming from a single pole⁸, strong obliquity between the rotation axis and the magnetic axis is still required if the pulsar is to have a normal magnetosphere. Therefore, either initially weak magnetic fields or actual field decay can be inferred from the observed low torque.

The rapid spin of the star may allow some new constraints to be put on the emission physics of radio pulsars. Even though the polar field lines in the emission region have an opening angle $\geq \sin^{-1}(R_\bullet/R_\text{L})^{1/2} \sim 20^\circ$, the formation of a sharp pulse is still possible in the polar beaming model of pulsar emission, if the emission arises in the much thinner boundary between the pair plasma and the neighbouring quasi-vacuum zones^{6,8,36}. Temporal resolution of the radio pulse would test whether this hypothesis is consistent with the geometry of the emission zone. In addition, aberration may strongly affect the arrival of a pulse at different frequencies, if different frequencies are emitted at different radii. Dedispersed pulsed observations with $10 \mu\text{s}$ time resolution should be able to detect this effect, if the emission zone is extended radially over distances $\Delta r \geq R_\bullet \sim 10 \text{ km}$, as is expected if the radius to frequency mapping interpretation of pulsars' radio spectra is correct³⁷.

The rapidly rotating millisecond pulsar PSR1937+214, therefore, provides a new testing ground, both for ideas about pulsar physics and for concepts of their evolution.

I thank D. Backer, J. Barnard, R. D. Blandford, Marc Davis, J. Grindlay, C. Heiles, S. Kulkarni, C. E. Max, C. F. McKee, S. Rappaport, R. Wagoner and S. D. M. White for informative and stimulating conversations.

This research was supported by NSF grant AST 79-23243, by California Space Institute grant CS 82-41, and also by the taxpayers of California.

Received 30 November 1982; accepted 27 January 1983.

1. Backer, D. C., Kulkarni, S., Heiles, C., Davis, M. & Goss, W. M. *Nature* **300**, 615 (1982).
2. Sturrock, P. A. *Astrophys. J.* **164**, 529 (1971).
3. Ruderman, M. A. & Sutherland, P. *Astrophys. J.* **196**, 51 (1975).
4. Arons, J. & Scharlemann, E. T. *Astrophys. J.* **231**, 854 (1979).
5. Barnard, J. & Arons, J. *Astrophys. J.* **254**, 713 (1982).
6. Arons, J. *Astrophys. J.* (in the press).
7. Michel, F. C. *Astrophys. J.* **158**, 727 (1969).
8. Arons, J. *Space Sci. Rev.* **24**, 437 (1979).
9. Michel, F. C. & Pellat, R. in *Proc. IAU Symp.* No. 95, 37 (1981).
10. Michel, F. C. *Rev. mod. Phys.* **54**, 1 (1982).
11. Pacini, F. *Nature* **216**, 567 (1967).
12. Ostriker, J. P. & Gunn, J. E. *Astrophys. J.* **157**, 1395 (1969).
13. Arons, J. in *Proc. Workshop on Electron-Positron Pairs in Astrophysics* (eds Burns, M., Harding, A. K. & Ramaty, R.) (American Institute of Physics, New York, in the press).

14. Max, C. E. & Perkins, F. *Phys. Rev. Lett.* **27**, 1342 (1972).
15. Arons, J., Norman, C. A. & Max, C. E. *Phys. Fluids* **20**, 1302 (1977).
16. Romeiras, F. J. *J. Plasma Phys.* **22**, 201 (1979).
17. Leboeuf, J. N. *et al. Phys. Rev. A* **25**, 1023 (1982).
18. Rees, M. J. & Gunn, J. E. *Mon. Not. R. astr. Soc.* **167**, 1 (1974).
19. Kennel, C. F. & Coroniti, F. V. Preprint (Univ. California, Los Angeles, 1983).
20. Djorgovski, S. *Nature* **300**, 618 (1982).
21. Oda, M. in *Proc. int. Summer School and Workshop on Plasma Astrophysics* (ed. Guyenne, T. D.) 251 (European Space Agency, Paris, 1981).
22. Wilson, A. S. *Mon. Not. R. astr. Soc.* **157**, 229 (1972).
23. Swinbank, E. & Pooley, G. *Mon. Not. R. astr. Soc.* **186**, 775 (1979).
24. Backer, D. C., Kulkarni, S. & Taylor, J. H. *Nature* **301**, 314 (1983).
25. Manchester, R. N. & Taylor, J. H. *Pulsars* (Freeman, San Francisco, 1977).
26. Phinney, E. S. & Blandford, R. D. *Mon. Not. R. astr. Soc.* **194**, 137 (1981).
27. Taylor, J. H. *Pap. at XIth "Texas" Conf. on Relativistic Astrophysics* (1982).

28. Smarr, L. & Blandford, R. D. *Astrophys. J.* **207**, 279 (1976).
 29. Alpar, A., Cheng, A. F., Ruderman, M. A. & Shaham, J. *Nature* **300**, 728 (1982).
 30. Arons, J. & Lea, S. M. *Astrophys. J.* **207**, 914 (1976).
 31. Arons, J. & Lea, S. M. *Astrophys. J.* **235**, 1016 (1980).
 32. Ghosh, P. & Lamb, F. K. *Astrophys. J.* **232**, 259 (1979).
 33. van den Heuvel, E. P. J. *Proc. IAU Symp.* No. 95, 379-396 (1981).

34. Rappaport, S. & Joss, P. in *Accretion Driven X-Ray Sources* (ed. Lewin, W. H.) (Cambridge University Press, in the press).
 35. Kippenhahn, R. & Weigert, A. Z. *Astrophys.* **65**, 251 (1976).
 36. Arons, J. in *Proc. int. Summer School and Workshop on Plasma Astrophysics*, (ed. Guyenne, T. D.) 273 (European Space Agency, Paris, 1981).
 37. Cordes, J. *Astrophys. J.* **222**, 1006 (1978).

Structure and expression of a cloned cDNA for human interleukin-2

Tadatsugu Taniguchi, Hiroshi Matsui, Takashi Fujita, Chikako Takaoka, Nobukazu Kashima*, Ryota Yoshimoto* & Junji Hamuro*

Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Toshima-ku, Tokyo 170, Japan

* Central Research Laboratories, Ajinomoto Co. Inc., Totsuka-ku, Yokohama 244, Japan

A cDNA coding for human interleukin-2 (IL-2) has been cloned from a cDNA library prepared from partially purified IL-2 mRNA. The DNA sequence codes for a polypeptide which consists of 153 amino acids including a putative signal sequence. A biologically active polypeptide, characteristic of human IL-2, was produced when the cDNA was fused to a simian virus 40 promoter sequence and used to transfect cultured monkey COS cells.

INTERLEUKIN-2 (IL-2), formerly referred to as T-cell growth factor, is a lymphokine which is produced by lectin- or antigen-activated T cells^{1,2}. The reported biological activities of IL-2 include stimulation of the long-term *in vitro* growth of activated T-cell clones^{1,3}, enhancement of thymocyte mitogenesis^{4,5} and induction of cytotoxic T-cell reactivity and plaque-forming cell responses in cultures of nude mouse spleen cells^{6,7}. As this lymphocyte regulatory molecule can be used to maintain culture of functional monoclonal T cells it has a key role in the study of the molecular nature of T-cell differentiation, the mechanism of differentiated T-cell functions as well as T-cell antigen receptors. In addition, like interferons (IFNs), IL-2 has been shown to augment natural killer cell activity⁸, suggesting a potential use in the treatment of neoplastic diseases.

The IL-2 molecule has been characterized physicochemically in several laboratories. Human IL-2 has a molecular weight (MW) of 15,000 (ref. 9). Molecular heterogeneity of human IL-2 described previously is at least partly due to the different extent of sialylation of the molecules¹⁰. mRNA encoding IL-2 has also been extracted from various cell sources and translated in *Xenopus laevis* oocytes and in rabbit reticulocyte cell-free systems^{9,11,12} but much less data are available with respect to the structure of IL-2.

Here we describe the first cloning and sequence analysis of a cDNA coding for human IL-2. Expression of the cDNA in cultured monkey COS cells gave rise to a protein product characteristic of authentic human IL-2.

Hybrid plasmids containing IL-2 cDNA

We prepared IL-2 mRNA from a human leukaemic T-cell line, designated Jurkat-111, cloned from Jurkat-FHCRC (ref. 13). The cells (3×10^6 cells per ml) were stimulated by concanavalin A (Con A, $25 \mu\text{g ml}^{-1}$) for IL-2 production and, 6 h later, polyadenylated RNA (poly(A) RNA) was prepared¹⁴ and fractionated by centrifugation on a sucrose density gradient¹⁵. IL-2 mRNA activity was monitored by injecting aliquots of each fraction into *X. laevis* oocytes¹⁶ and assaying the incubation medium for the ability to stimulate the incorporation of tritiated thymidine (^3H -TdR) by an IL-2-dependent T-cell line (CTLL-2) according to the methods described by Gillis *et al.*².

One peak of IL-2 mRNA activity was constantly observed which sedimented at $\sim 11.5\text{S}$ (Fig. 1). As has been reported by Efrat *et al.*¹², we also occasionally found the mRNA activity which corresponds to 13S. We believe that this 13S mRNA in our RNA preparation is an aggregated form of the 11.5S mRNA (see discussion). RNA from the most active fraction was used as template to synthesize double-stranded cDNA by the standard procedure^{17,18}. The cDNA was size-fractionated and

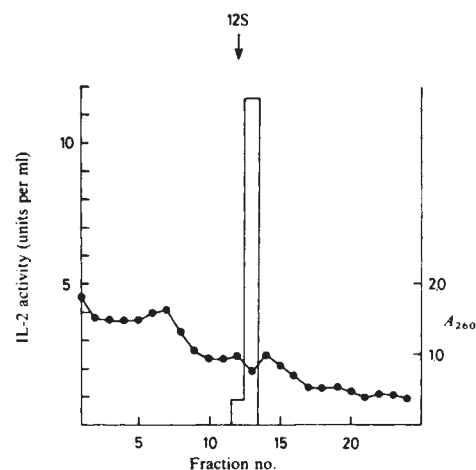


Fig. 1 Sedimentation analysis of human IL-2 mRNA. Human leukaemic T cells (Jurkat-111) were grown in RPMI 1640 medium containing 10% FCS in roller bottles (volume 1:1) to the final cell density of 1.5×10^6 cells per ml. Cells, concentrated to 3×10^6 cells per ml in the same medium containing 1% FCS, $100 \mu\text{g ml}^{-1}$ each of penicillin and streptomycin, were then induced for IL-2 production with Con A ($25 \mu\text{g ml}^{-1}$). In these conditions, IL-2 production in culture medium increased to 2,000 units per ml. Total RNA was extracted from the induced cells and poly(A)-containing RNA was purified by oligo(dT)-cellulose chromatography¹⁴. The mRNA (2.4 mg) was heated in 0.5 mM EDTA at 70°C for 2 min and fractionated on a 5–25% linear sucrose gradient in 50 mM Tris-HCl, pH 7.5, 0.2 M NaCl and 1 mM EDTA by centrifugation at 26,000 r.p.m. for 26 h at 4°C in a Beckman SW28 rotor. ^{32}P -labelled *Pst*I-*Hind*III fragments of pBR322 which sediment at 12S and 9S were included as size markers. RNA from individual fractions was ethanol precipitated, dissolved in water and each aliquot was assayed for IL-2 mRNA by injection in *X. laevis* oocytes¹⁶ at a concentration of $0.3 \mu\text{g ml}^{-1}$. After 24 h incubation at 24°C , the media were assayed for IL-2 activity by measuring ^3H -TdR incorporation of cloned cytotoxic T cells (CTLL-2) essentially as described previously². Briefly, 4,000 CTLL-2 cells were seeded in $100 \mu\text{l}$ of RPMI 1640 medium containing 2% FCS in 96-well flat-bottom microplates together with $100 \mu\text{l}$ of the serially diluted translation products. After 20 h at 37°C , cells were pulsed for 4 h with $0.5 \mu\text{Ci}$ of ^3H -TdR, collected onto glass fibre strips with the aid of automated cell harvester and then the incorporated radioactivity was determined. One unit per ml of IL-2 stimulated 50% of the maximum ^3H -TdR incorporation which was observed with standard rat IL-2 preparation.

material longer than 600 base pairs (bp) was inserted into pBR322 by the standard G-C tailing method¹⁹. The resulting hybrid plasmid was used to transform *Escherichia coli* K-12

strain χ 1776 (ref. 20) to prepare a cDNA library as described previously¹⁸.

To identify an IL-2 cDNA clone, we screened the cDNA library by a mRNA hybridization translation assay^{18,21-23}. Hybrid plasmids were prepared from groups of 24 bacterial clones, and about 50 μ g of each plasmid group were cleaved with *Hind*III, denatured and bound to nitrocellulose filters. The filters were hybridized with 400 μ g of poly(A) RNA from induced Jurkat-111 cells. Hybridized RNA was recovered from the filters, purified and injected into *Xenopus* oocytes to determine the IL-2 mRNA activity. Control hybridization was carried out with pBR322.

One out of 18 groups, each consisting of 24 clones, gave a positive result (group 3) by this assay while others were clearly negative (Table 1). We then prepared plasmid DNAs separately from single clones from group 3 and assayed them for the presence of IL-2 mRNA sequences as above. As shown in Table 1, only plasmid DNA from clone 16, designated p3-16, gave a positive response.

Plasmid p3-16 contained a cDNA insert consisting of about 650 bp, which is apparently shorter than the corresponding 11.5S mRNA. We therefore prepared another cDNA library according to the procedure of Land *et al.*²⁴ and screened it for cDNA clones containing larger inserts by *in situ* hybridization²⁵ with ³²P-labelled p3-16 cDNA insert. By this screening, we identified a clone containing a plasmid, pIL2-50A, whose cDNA insert consists of about 880 bp.

Jurkat cells do not produce IL-2 without a stimulus such as Con A (ref. 13). Accordingly, we never detected IL-2 mRNA activity in the poly(A) RNA from mock-induced Jurkat-111 cells (unpublished observation). We examined the presence and the size of the mRNA which correspond to pIL2-50A cDNA insert by means of RNA blot analysis²⁶ in poly(A) RNA isolated from induced or mock-induced Jurkat cells. As shown in Fig. 2, the cloned cDNA hybridized only to the mRNA from induced cells that was of very similar size to IL-2 mRNA.

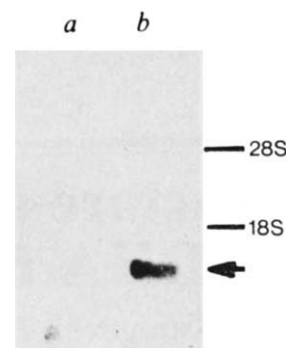


Fig. 2 Analysis of the mRNA which corresponds to pIL2-50A cDNA insert. Poly(A) RNA from Con A-induced or mock-induced Jurkat-111 cells was prepared as described in Fig. 1 legend. 10 μ g of the RNA was denatured and subjected to RNA blotting analysis as described by Thomas²⁶ using nick-translated pIL2-50A cDNA as the probe. *a*, Poly(A) RNA from mock-induced Jurkat cells; *b*, poly(A) RNA from Con A-induced Jurkat cells. Ribosomal RNA (28S and 18S) was run separately as size markers. The arrow indicates the position of the positive band which appeared when poly(A) RNA containing 12S human IFN- β ₁ mRNA was analysed in parallel using nick-translated IFN- β ₁ cDNA^{18,35} as the probe.

Nucleotide sequence of the cDNA insert

The complete nucleotide sequence of the *Pst*I insert from pIL2-50A was determined by the procedure of Maxam and Gilbert²⁷. Restriction endonuclease cleavage map of the cDNA insert and the sequencing strategy are illustrated in Fig. 3*a*. The DNA sequence of the insert contains a single large open reading frame (Fig. 3*b*). The first ATG, which usually serves as the initiation codon in eukaryotes²⁸, is found at nucleotides

Table 1 mRNA hybridization translation assay for the detection of IL-2 cDNA from pools of transformed *E. coli*

	DNA samples	IL-2 activity (units per ml)	DNA samples	IL-2 activity (units per ml)	DNA samples	IL-2 activity (units per ml)
First screening: groups of 24 clones	1	<0.1	7	<0.1	13	<0.1
	2	<0.1	8	<0.1	14	<0.1
	3	4.2	9	<0.1	15	<0.1
	4	<0.1	10	<0.1	16	<0.1
	5	<0.1	11	<0.1	17	<0.1
	6	<0.1	12	<0.1	18	<0.1
Second screening: single clones from group 3	1	<0.1	9	<0.1	17	<0.1
	2	<0.1	10	<0.1	18	<0.1
	3	<0.1	11	<0.1	19	<0.1
	4	<0.1	12	<0.1	20	<0.1
	5	<0.1	13	<0.1	21	<0.1
	6	<0.1	14	<0.1	22	<0.1
	7	<0.1	15	<0.1	23	<0.1
	8	<0.1	16	16.0	24	<0.1

Hybrid plasmid DNA containing human IL-2 cDNA was constructed as follows. Double-stranded cDNA was prepared by standard procedures^{17,18} using 18 μ g of mRNA from fraction 13 (Fig. 1). The cDNA of over 600 bp (2.4 μ g) was obtained after fractionation on a sucrose density gradient. The cDNA was then extended with dCMP residues using terminal deoxynucleotidyl transferase⁴¹ and an aliquot (50 ng) was annealed with 250 ng of dGMP-elongated, *Pst*I-cleaved pBR322 (ref. 41). The resulting hybrid plasmids were used to transform *E. coli* χ 1776 (ref. 19). From this library (total of 2,000 clones), 432 clones were picked up arbitrarily and individually inoculated into wells of microtitre plates containing 200 μ l of L broth⁴² and 10 μ g ml⁻¹ tetracycline. The plates were incubated at 37 °C for 16 h. To carry out the hybridization translation assay, 100 μ l of each bacterial culture from pools of 24 clones was mixed and the resulting mixture was used to inoculate 250-ml cultures from which plasmid DNA was prepared as described previously⁴². Thus, a total of 18 plasmid DNA corresponding to 432 clones were prepared. Each plasmid (50 μ g) was cleaved with *Hind*III and bound individually to a nitrocellulose filter (Schleicher & Schüll, 2.5 cm diameter, 0.45- μ m pores) essentially as described by Harpold *et al.*²⁰. Filters were then hybridized with 400 μ g poly(A) RNA from induced Jurkat cells in 600 μ l the hybridization solution containing 50% (v/v) formamide, 20 mM PIPES at pH 6.4, 0.4 M NaCl, 2 mM EDTA and 0.1% SDS. After 18 h at 37 °C, filters were washed three times with 200 ml of 10 mM PIPES, pH 6.4, 0.15 M NaCl, 1 mM EDTA and 0.2% SDS at 65 °C and three times with 50 ml of 1 mM PIPES, pH 6.4, and 10 mM NaCl at 25 °C. The filter-hybridized RNA was recovered by heating the filter three times in 0.6 ml of 0.5 mM EDTA and 0.1% SDS at 95 °C for 1 min and the RNA was purified by oligo(dT)-cellulose chromatography. The RNA sample was finally dissolved in 2.5 μ l of 10 mM Tris-HCl (pH 7.5) and 88 mM NaCl and injected into 10–15 *Xenopus laevis* oocytes (50 nl per oocyte). For the second screening, the assay was carried out essentially as above except that 25 μ g of plasmid DNA from individual clones was bound to the nitrocellulose filter of 1 cm diameter (binding efficiency was about 75%). IL-2 activity in oocyte medium was determined as described in Fig. 1 legend. No IL-2 mRNA activity was detectable (0.1 units ml⁻¹) when pBR322 DNA was used for the assay.

48–50 from the 5' end. This ATG is followed by 152 codons before the termination triplet TGA is found at nucleotides 507–509. A stretch of A residues corresponding to the 3' poly(A) terminus of the mRNA is found at the end of the cDNA and this is preceded by the hexanucleotide AATAAA (position 771–776) which is usually found in most eukaryotic mRNAs²⁹. To examine whether the 5'-terminal region of the mRNA is copied efficiently in the cDNA of pIL2-50A, a single-stranded *RsaI*-*DdeI* fragment (position 52–84), ³²P-labelled at the 5' *DdeI* site, was prepared and used to prime the reverse transcription of 11.5S IL-2 mRNA into single-stranded ³²P-cDNA. The size of this ³²P-cDNA product was 85–87 nucleotides (unpublished results), indicating that the cDNA insert of pIL2-50A covers the length of the IL-2 mRNA except perhaps for a few nucleotides.

A primary structure of the human IL-2 polypeptide consisting of 153 amino acids could be deduced as shown in Fig. 3b and its MW was calculated to be 17,631.7. As has been reported as a common feature in most of the secretion proteins known to date³⁰, the N-terminal region of the deduced IL-2 polypeptide sequence is also quite hydrophobic and this region probably serves as a signal peptide which is cleaved off in the secretion process of mature IL-2. Such cleavage could occur between Ser and Ala at positions 20 and 21, respectively, as similar cleavage sites have often been found in other secretion proteins³⁰. If so, mature human IL-2 would contain 133 amino acids with the calculated MW of 15,420.5; this is comparable with the reported MW of human IL-2 protein from Jurkat cells (15,000)⁹, but cleavage could take place at an alternative position somewhere around this site. Human IL-2 from Jurkat cells is reportedly non-glycosylated *in vivo*¹⁰. Consistent with this observation is the absence of potential *N*-glycosylation sites³¹ (Asn-X-Ser or Asn-X-Thr) in the deduced sequence of human IL-2. However, the presence of modifications other than *N*-glycosylation, such as *O*-glycosylation and sialylation, cannot be ruled out^{10,32}. The cDNA sequence also predicts that the IL-2 molecule has a neutral isoelectric point, having an equal number of basic (Arg + Lys) and acidic (Asp + Glu) amino acids and this is consistent with the value obtained with partially purified IL-2 by Gillis *et al.*⁹ and by Ruscetti and Gallo³³

(*pI* = 6.8–7.1). When the structure of IL-2 polypeptide was compared with those of other soluble factors such as interferons and growth hormones to which IL-2 could have some functional relationships, little homology was observed. With respect to human IFN- γ , another lymphokine, homology covering three consecutive amino acids occurs at amino acids 61–63 (Lys-Asn-Phe)³⁴ which are also found in human IL-2 at positions 96–98. Homologous amino acid sequences can also be found between human IL-2 (Leu-Lys-Pro-Leu-Glu-Glu-Val-Leu at positions 83–90) and human IFN- β_1 (Leu-Lys-Thr-Val-Leu-Glu-Glu-Lys-Leu at positions 98–106)³⁵. Further detailed analysis may be required to search for other homologies with respect to their secondary structure.

Expression of IL-2 cDNA in monkey cells

Despite the extensive characterization of human IL-2 molecules in several laboratories, nothing is known about the primary structure of this protein. It is therefore essential to prove that the cDNA sequence of pIL2-50A actually codes for IL-2 by obtaining expression of this cloned gene.

We used a mammalian cell system to express the cDNA because, unlike bacteria such as *E. coli*, it should be possible to obtain the protein product being properly processed and secreted in the culture medium³⁴.

A plasmid which should direct the synthesis of human IL-2 in mammalian cells was constructed as follows. A plasmid pCE-1 was constructed from pKCR (ref. 36) and pBR328 (ref. 37) by a series of modification procedures as illustrated in Fig. 4. This plasmid contains a single *PstI* site just downstream of the SV40 early promoter and upstream of the part of the rabbit β -globin chromosomal gene containing one intron. The plasmid also contains the replication origin of SV40 as well as the polyadenylation site for the early gene³⁶. The cDNA insert of pIL2-50A was isolated by *PstI* cleavage and inserted into the *PstI* site of pCE-1. Thus a plasmid pCEIL-2, in which the IL-2 structural gene should be transcribed from the early promoter of SV40 in appropriate host cells, was obtained (Fig. 4).

This plasmid was digested by *HhaI* and then introduced by DNA transfection into the transformed monkey cell line COS-7 which allows replication of DNA containing SV40 origin sequences³⁸. It seems important to digest the plasmid with *HhaI* before transfection for the efficient expression of cDNA because sequences which could hamper replication of the transfected DNA in COS cells³⁹ can be removed from the essential part of the plasmid for cDNA expression by this procedure. In our previous experiments, the expression level went down approximately 5–10-fold without this procedure when human IFN- β_1 and γ cDNA were expressed similarly (unpublished results).

Two to three days after transfection, the cultured cell medium were assayed for human IL-2 activity. As shown in Table 2, media from pCEIL-2 transfected cultures contained IL-2 activity. No IL-2 activity was detectable in the media from cells transfected with pCE-1 DNA.

IL-2 activity from monkey cells

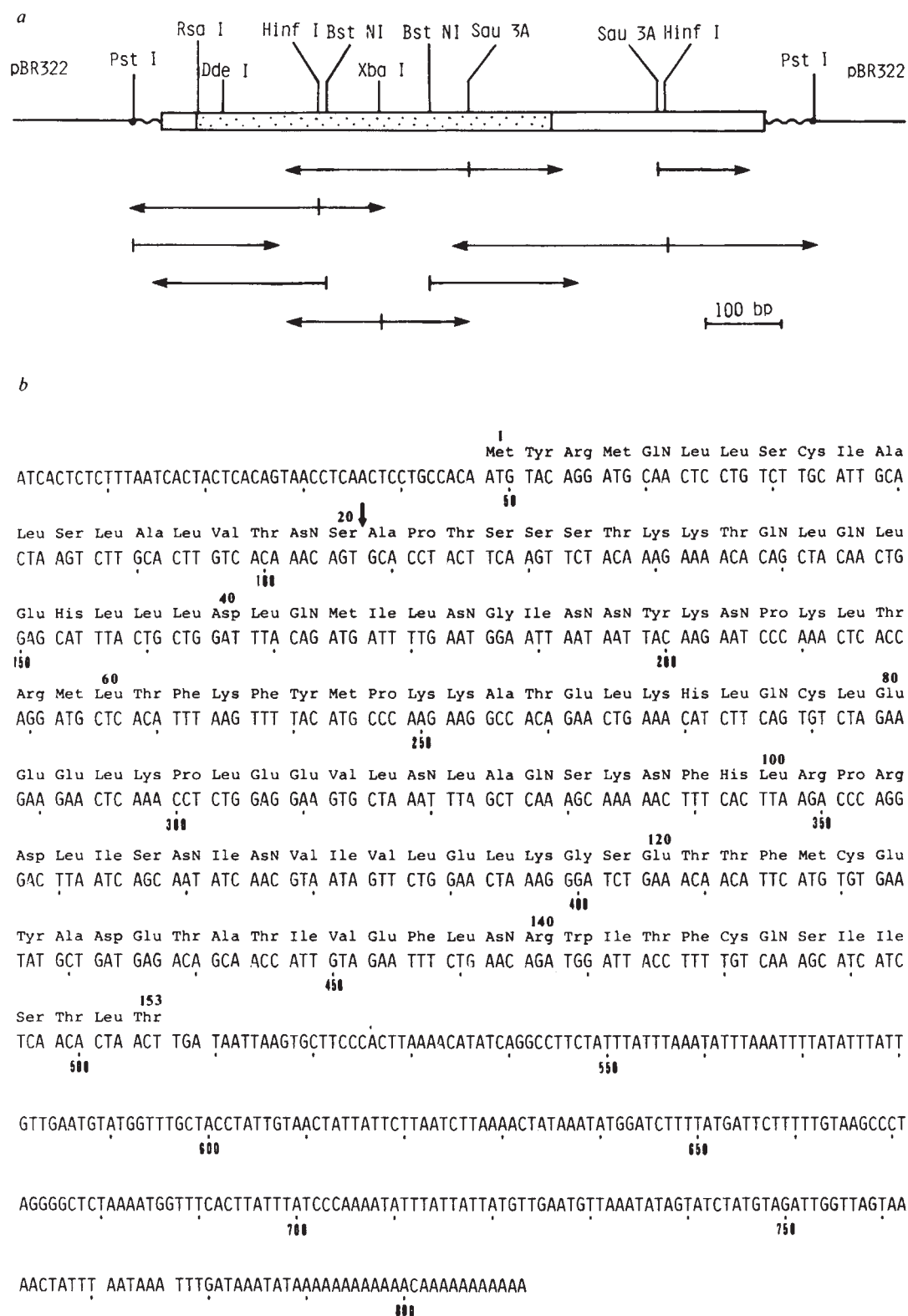
Properties of the IL-2 activity secreted from pCEIL-2 transfected COS cell were examined. As shown in Table 2, it stimulates the ³H-TdR incorporation by the IL-2-dependent CTLL-2 cell line and also stimulates *in vitro* T-cell growth. This activity was neutralized by a monoclonal antibody raised against purified human IL-2 in a BALB/c mouse (unpublished results). The dose-response curve of this IL-2 activity was comparable to that of native IL-2 as shown in Fig. 5. Sensitivity of this activity to trypsin treatment indicated that a protein is responsible for the activity (results not shown). In addition, Sephadex G-100 column chromatography showed that the MW of the IL-2 produced in COS cells was ~15,000 and was indistinguishable from authentic human IL-2 generated by Con A-stimulated Jurkat cells (result not shown). The protein encoded by the cloned cDNA sequence of pIL2-50A therefore shows many of the known properties of human IL-2.

Table 2 Detection of the IL-2 activity in the culture medium of COS-7 cells transfected by pCEIL-2

Transfected DNA	Medium collected after transfection (h)	IL-2 activity measured by	
		³ H-TdR incorporation (units per ml)	T-cell growth
pCEIL-2	48	1.5	++
		1.5	++
	72	14.0	++++
		6.0	++++
pCE-1	48	<0.1	—
		<0.1	—
	72	<0.1	—
		<0.1	—

Fresh monolayers of COS-7 cells in wells of 1.5 cm diameter with 1 ml of Dulbecco's minimal essential medium (DMEM, Gibco) were transfected with 1 μ g of *HhaI*-digested pCEIL-2 or pCE-1 DNA by using the calcium phosphate co-precipitation technique essentially as described previously⁴³. After 15 h at 37 °C, medium were replaced with 1 ml of fresh DMEM containing 8% fetal calf serum (FCS). Each medium, collected at appropriate intervals (indicated in the table) was kept at 4 °C until use. The IL-2 assay by ³H-TdR incorporation was performed as described in Fig. 1 legend. To measure the ability to stimulate *in vitro* T-cell growth⁴⁴, five CTLL-2 cells were mixed with the undiluted culture medium to be assayed for IL-2 (0.1 ml) and 0.1 ml of DMEM containing 2% FCS. Cells were cultured for 72 h in 96-well flat bottom microplate and cell number was counted by microscopy: —, <1 viable cell; ++, 5–20 cells; +++, 50–100 cells.

Fig. 3 a, Restriction endonuclease cleavage maps of the cDNA insert of pIL2-50A. To obtain pIL2-50A, another cDNA library was prepared essentially according to the procedure of Land *et al.*²⁴. Briefly, 1.6 µg of single-stranded cDNA was synthesized by using 4 µg of mRNA from fraction 13 (Fig. 1), elongated by dCMP residues, and double-stranded cDNA was synthesized by using oligo(dG)₁₂₋₁₈ as the primer for DNA polymerase I (Klenow fragment). The cDNA (0.6 µg) longer than the 680-bp DNA size marker was obtained by sucrose gradient centrifugation and inserted into the *Pst*I site of pBR322 by the standard G-C tailing method⁴¹. After transformation of *E. coli* χ 1776 by the hybrid plasmids, ~2,500 colonies were screened by *in situ* hybridization²⁵ with nick-translated p3-16 cDNA insert as the probe and the colony containing pIL2-50A was identified. The rectangle represents the cDNA insert, and the dotted area the protein coding region. Wavy lines indicate G-C junctions. The strategy of the DNA sequence analysis is presented in the lower part of the figure. Arrows indicate the extent of sequencing of each segment analysed without ambiguity. **b**, Nucleotide sequence and deduced amino acid sequence of the plasmid pIL2-50A cDNA insert. The entire sequence was determined by the chemical method of Maxam and Gilbert²⁷. The arrow indicates the conjectured cleavage site by a signal peptidase³⁰. Numbers above each line refer to amino acid position and numbers below each line to nucleotide position.



Discussion

A cDNA library was prepared from the fractionated poly(A) mRNA which sedimented at around 11.5S and showed the highest IL-2 mRNA activity. From this library, a hybrid plasmid, p3-16, containing the cDNA copy of human IL-2 was identified by an IL-2 mRNA hybridization translation assay. Subsequent rescreening of the library with the ³²P-cDNA probe indicated that the frequency of the IL-2-specific colonies in the cDNA library was only 1:2,000. This means that the frequency of the IL-2-specific mRNA in our poly(A) mRNA fraction is

approximately 1:20,000 as about 10-fold enrichment of the IL-2 mRNA is achieved by the sucrose density gradient fractionation¹⁸. As has been reported by Efrat *et al.*¹², we also occasionally found the IL-2 mRNA peak at 13S. However, when the 13S mRNA was subjected to RNA blot analysis with the same probe as above, only one positive band which co-migrates with 11.5S IL-2 mRNA was detected (data not shown). We conclude that the 13S IL-2 mRNA in our mRNA preparation was an aggregated form of 11.5S mRNA. A similar observation has been made by Gray *et al.* with human IFN- γ mRNA³⁴.

Fig. 4 Plasmid DNA constructed for the expression of IL-2 cDNA in monkey COS-7 cells. The plasmid pKCR (ref. 36) consists of: (1) Segments of SV40 DNA (shown as hatched blocks) containing the early gene promoter and an origin of replication (0.725–0.648 map units) and a polyadenylation site from the early gene (0.169–0.144 map units). (2) A part of the rabbit β -globin gene (shown as open blocks) (*Bam*HI–*Pvu*II fragment). (3) A segment from pBR (*Eco*RI–*Bam*HI fragment) containing an origin of replication and ampicillin resistance gene. This plasmid was cleaved by *Bam*HI and, after filling both ends by DNA polymerase I (Klenow fragment), a synthetic *Pst*I linker DNA was introduced to construct pKCR(*Pst*I). Plasmid pKCR(*Pst*I) was cleaved by *Sal*I, treated by the Klenow fragment to fill the ends and then partially cleaved by *Eco*RI to obtain *Eco*RI–*Sal*I fragment which contains the whole DNA derived from SV40 and the globin gene. This fragment was then ligated to a piece of pBR328 DNA which contains the tetracycline resistance gene and an origin of replication³⁷ as outlined. The resulting plasmid pCE-1 contains a single *Pst*I site just downstream of the SV40 early promoter. cDNA insert of pIL2-50A was excised by *Pst*I cleavage and ligated to *Pst*I-cleaved pCE-1 to construct pCEIL-2 in which expression of the IL-2 structural gene should be under control of the SV40 early promoter. Plasmid pCE-1 was originally constructed for the cDNA cloning by the G-C tailing method⁴¹ in bacteria and direct expression in mammalian cells.

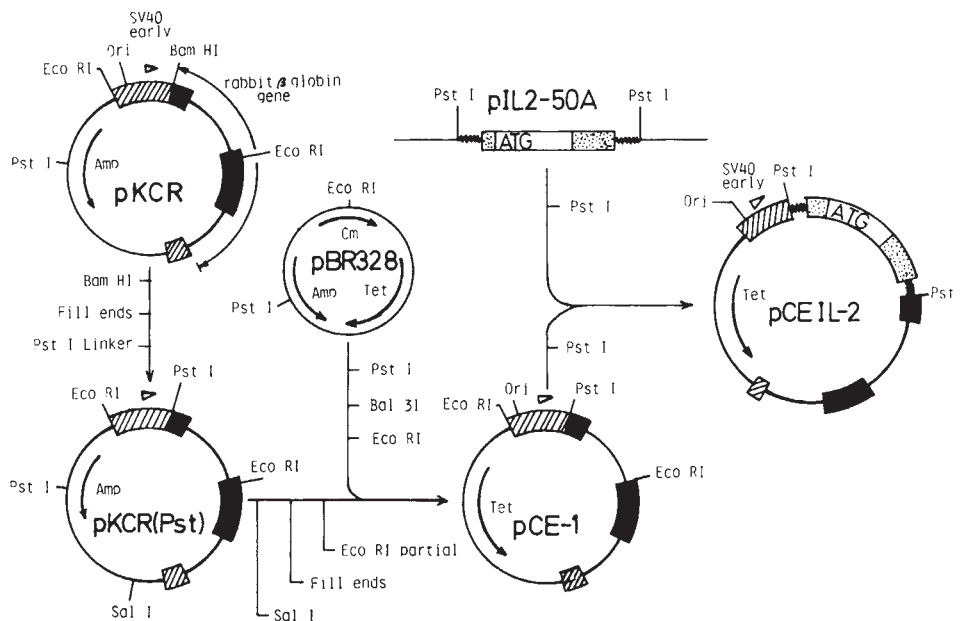
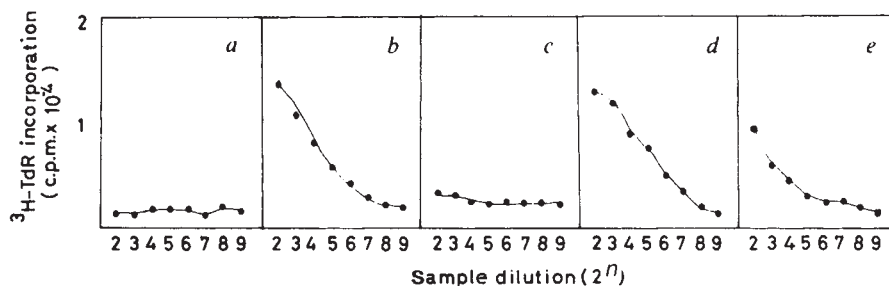


Fig. 5 Dose-response curve for the stimulation of 3 H-TdR incorporation by an authentic and cDNA-directed IL-2 preparation. The IL-2-dependent 3 H-TdR incorporation by the CTLL-2 cells has been measured as described in Fig. 1 legend. *a*, Medium alone which was used for COS cell expression. *b*, Cell medium collected 72 h after transfection of pCEIL-2 DNA to COS-7 cells (determined to be 16 units per ml). *c*, Cell media collected 72 h after transfection of pCE-1 DNA to COS-7 cells. *d*, IL-2 produced by Jurkat-111 cells (25 units per ml). *e*, IL-2 produced by Jurkat-111 cells (5 units per ml).



Sequence analysis of the cloned cDNA showed that the cDNA encodes a polypeptide of 153 amino acids, 20 of which could possibly be cleaved off during the secretion of mature IL-2. The identity of the deduced amino acid sequence has mostly been confirmed with a separate cDNA clone (unpublished observation). The deduced amino acid sequence contained no potential *N*-glycosylation site. Robb and Smith¹⁰ reported that IL-2 prepared from human tonsil cells were variably modified, in contrast to the preparation from Jurkat cells. It remains to be seen whether this is due to the difference in the producer cell type or in the gene structure.

To prove ultimately that the cloned cDNA corresponds to that of IL-2, we have expressed it in monkey COS cells. The expression product manifested biological activities characteristic of human IL-2. The activity found in COS cell media (the maximum value we have obtained so far is 70 units per ml) could become higher by removing the G-C stretch which is present between the SV40 promoter and the cDNA as reported

by Simonsen *et al.* in the expression of human IFN- γ cDNA⁴⁰. Expression of the IL-2 cDNA in *E. coli* is in progress and it will soon become possible by the use of cloned cDNA to produce this immunoregulatory molecule in large quantity for various purposes.

The cloning and analysis of the human IL-2 cDNA raise many questions. How many copies of IL-2 genes are present in the human genome? How is the gene organized in the chromosome? Is the gene linked to other lymphokine genes? Our preliminary analysis of the human genomic DNA indicates that the IL-2 gene exists in a single copy. Further studies will provide answers to these questions.

We thank Drs R. Breathnach and Y. Gulzman for providing us with pKCR and COS-7 cells, respectively; Dr H. Sugano for his continuous support and advice; Drs M. Muramatsu, T. Akashi, T. Miyata, K. Sano and S. Nagata for their support and valuable discussions and Dr Y. Fujii-Kuriyama for critical reading of the manuscript.

Received 13 January; accepted 10 February 1983.

- Morgan, D. A., Ruscetti, F. W. & Gallo, R. *Science* **193**, 1003–1007 (1976).
- Gillis, S., Ferm, M. M., Ou, W. & Smith, K. *J. Immunology* **120**, 2027–2033 (1978).
- Gillis, S. & Smith, K. A. *Nature* **268**, 154–156 (1977).
- Chen, B. M. & DiSabato, G. *Cell. Immun.* **22**, 211–224 (1967).
- Shaw, J., Monticone, V. & Paetkau, V. *J. Immunology* **120**, 1967–1973 (1978).
- Gillis, S., Baker, P. E., Union, N. A. & Smith, K. A. *J. exp. Med.* **149**, 1960–1968 (1979).
- Wagner, H., Hardt, C., Heek, K., Rollinghoff, M. & Pfizenmaier, K. *Nature* **284**, 278–280 (1980).

- Henney, C. S., Kuribayashi, K., Kern, D. E. & Gillis, S. *Nature* **291**, 335–338 (1981).
- Gillis, S. *et al. Immun. Rev.* **63**, 167–209 (1982).
- Robb, R. J. & Smith, K. A. *Molec. Immun.* **18**, 1087–1094 (1981).
- Bleackley, R. C. *et al. J. Immun.* **127**, 2432–2435 (1981).
- Efrat, S., Pilo, S. & Kaempfer, R. *Nature* **297**, 236–239 (1982).
- Hansen, J. A., Martin, P. J. & Nowinski, R. C. *Immunogenetics* **10**, 247–252 (1980).
- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412 (1972).
- Taniguchi, T., Pang, R. H. L., Yip, Y. K., Henriksen, D. & Vilcek, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3469–3472 (1981).
- Gurdon, J. B., Lane, C. D., Woodland, H. R. & Marbaix, G. *Nature* **233**, 177–182 (1971).

17. Wickens, M. P., Buell, G. N. & Shimke, R. T. *J. biol. Chem.* **253**, 2483–2495 (1978).
18. Taniguchi, T. *et al. Proc. Japan Acad.* **553**, 464–469 (1979).
19. Villa-Komaroff *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3727–3731 (1978).
20. Curtiss, R. III *et al. Miami Winter Symp.* **13**, 99–114 (1977).
21. Harpold, M. M., Dobner, P. R., Evans, R. M. & Bancroft, F. C. *Nucleic Acids Res.* **5**, 2039–2053 (1978).
22. Nagata, S. *et al. Nature* **284**, 316–320 (1980).
23. Parnes *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2253–2257 (1981).
24. Land, H., Grez, N., Hansen, H., Lindermaier, W. & Schuetz, G. *Nucleic Acids Res.* **9**, 2251–2266 (1981).
25. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2330–2334 (1975).
26. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).
27. Maxam, A. W. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
28. Kozak, M. *Cell* **15**, 1109–1123 (1978).
29. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211–214 (1976).
30. Blobel, G. *et al. Symp. Soc. exp. Med.* **33**, 9–36 (1979).
31. Winzler, R. J. in *Hormonal Proteins and Peptides* (ed. Li, C. H.) 1–15 (Academic, New York, 1973).
32. Tomita, M. & Marchesi, V. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2964–2968 (1975).
33. Ruscetti, R. W. & Gallo, R. C. *Blood* **57**, 379–385 (1981).
34. Gray, P. W. *et al. Nature* **295**, 503–508 (1982).
35. Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y. & Muramatsu, M. *Gene* **10**, 11–15 (1980).
36. O'Hare, K., Benoist, C. & Breathnach, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1527–1531 (1981).
37. Soberon, X., Covarrubias, L. & Bolivar, F. *Gene* **9**, 287–305 (1980).
38. Gluzman, Y. *Cell* **23**, 175–182 (1981).
39. Lusky, M. & Botchan, M. *Nature* **293**, 79–81 (1981).
40. Simonsen, C. C. *et al. UCLA Symp.* **25**, 1–14 (1982).
41. Chang, A. C. Y. *et al. Nature* **275**, 617–624 (1978).
42. Bolivar, F. *et al. Gene* **2**, 95–113 (1977).
43. Frost, E. & Williams, J. *Virology* **91**, 39–50 (1978).
44. Stull, D. & Gillis, S. J. *Immun.* **126**, 1680–1683 (1981).

Gene and protein structure of a β -crystallin polypeptide in murine lens: relationship of exons and structural motifs

George Inana, Joram Piatigorsky, Barbara Norman, Christine Slingsby & Tom Blundell

Laboratory of Molecular and Developmental Biology, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20205, USA and Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, UK

A 23,000 molecular weight β -crystallin (β 23) of the murine eye lens is encoded in a 4.1 ± 0.3 -kilobase gene containing three introns. Each of the four exons seems to code for a separate structural motif of the protein, whose tertiary structure was predicted by an interactive computer graphics technique based on the crystallographic structure of bovine γ II-crystallin. The first exon also encodes a hydrophobic N-terminal peptide resembling membrane anchor sequences of other proteins. Our results indicate structural homology among the β - and γ -crystallin polypeptides, and link gene structure with protein structure in this superfamily of lens proteins.

THE transparent lens of vertebrates consists of anterior epithelial cells overlying posterior fibre cells; both the epithelial and fibre cells contain high concentrations of structural proteins called crystallins^{1,2}. There are four classes of lens crystallins (α -, β -, γ - and δ -crystallin). δ -Crystallin differs from the others by being confined to birds and reptiles^{3,4}. Each crystallin class is immunologically distinct and consists of a family of related polypeptides. Except for the γ -crystallins and one β -crystallin polypeptide (β s), which remain as monomeric proteins, the polypeptides of each crystallin class associate in various combinations to form higher molecular weight polymeric proteins. The crystallins are differentially synthesized during lens development and therefore are not distributed uniformly throughout the lens^{5,6}.

The crystallins are highly conserved proteins². For example, the primary structure of the α A₂-crystallin polypeptide has evolved at a rate of 2–3 amino acid substitutions per 100 residues per 100 Myr^{7,8}, which is comparable with the rate for cytochrome *c* and considerably slower than that for globin⁹. Hybridization experiments using mRNA and cloned genomic DNA¹⁰, protein and peptide analyses^{4,11} and spectroscopic tests^{12,13} indicate that δ -crystallin also has been structurally conserved throughout the evolution of birds and reptiles.

The basis for the apparent heterogeneity among the lens crystallins is unknown. Common structural features for the crystallins have been revealed recently, however, indicating homologies within this diverse group of proteins. Sequence determinations of the principal bovine β -crystallin polypeptide (β Bp, molecular weight (MW) $\sim$ 23K)¹⁴, a major murine β -crystallin polypeptide (β 23, MW $\sim$ 23K)¹⁵ and the predominant bovine γ -crystallin (fraction II) polypeptide (γ II, MW $\sim$ 20K)¹⁶ suggest that the β - and γ -crystallins have evolved from a common ancestral protein and belong to a superfamily¹⁷ of β γ -crystallins. Bovine γ II has been shown by X-ray analysis to have a two-domain β -sheet structure with four similar 'Greek key' motifs¹⁸. The primary structure of the N-terminal half of bovine

γ II¹⁸, bovine β Bp¹⁴ and murine β 23¹⁵ is partially homologous to that of the C-terminal half within each polypeptide, suggesting an intragenic duplication in the ancestral gene(s) of these proteins. The homology indicates that the proteins may have a common secondary structure and tertiary fold¹⁹. In the present investigation we have completed¹⁵ the deduced amino acid sequence of murine β 23 by analysis of its gene, constructed a three-dimensional model for β 23 by use of interactive computer graphics based on the crystallographic structure of bovine γ II (refs 18, 19) and related the putative structure of β 23 to the structure of its gene.

The murine β 23 gene

A cDNA (pM β 23Cr1, previously called pM β Cr1) containing the sequence for the mRNA of murine β 23, which has been cloned earlier in the bacterial plasmid pBR322 (ref. 15), was used here to isolate the β 23 gene. pM β 23Cr1 has been sequenced and contains 769 base pairs (bp) (including 100 poly(A) residues) of the β 23 mRNA, which is $\sim$ 1 kb long. The cDNA insert in pM β 23Cr1 encodes 183 residues of the 23K β -crystallin polypeptide and contains 120 bp of 3' untranslated non-poly(A) sequences in the mRNA. The 5' untranslated sequences and translational initiation site of the β 23 mRNA seem to be absent from pM β 23Cr1.

A Southern blot²⁰ showed that *Eco*RI-digested genomic DNA of the mouse has two DNA bands of $\sim$ 3.2 and 2.3 kb that hybridize to ³²P-pM β 23Cr1 (Fig. 1a). The lack of cross-hybridization to other murine β -crystallin gene sequences in these stringent conditions (see Fig. 1 legend) is consistent with similar results in the chicken where four different β -crystallin cDNA clones hybridized only to their respective gene sequences²¹. A genomic clone (λ M β 23Cr2) isolated from a DNA library (made in λ Charon 4A bacteriophage and containing embryonic BALB/c mouse chromosomal DNA partially digested with *Hae*III) also contained the 3.2- and 2.3-kb *Eco*RI DNA fragments that hybridize to ³²P-pM β 23Cr1 (Fig. 1b),

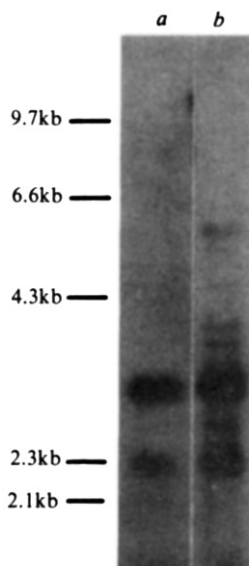


Fig. 1 Southern blots of *Eco*RI-digested mouse genomic DNA (a) and λ M β 23Cr2 (b) hybridized with 32 P-pM β 23Cr1. a, High molecular weight chromosomal DNA was isolated from mouse embryos, digested at 37 °C overnight to completion with 4 units of *Eco*RI (BRL) per μ g DNA, electrophoresed in a 0.8% agarose gel at 40 mA ($\sim 20 \mu$ g DNA per lane) and transferred to a nitrocellulose filter (HAHY; Millipore Corporation) by the Southern blotting technique²⁰. The blot was hybridized to murine 23K β -crystallin cDNA (pM β 23Cr1)¹⁵ labelled with 32 P by nick-translation⁵³ to $50\text{--}100 \times 10^6$ c.p.m. per μ g DNA. The hybridized, washed blot (0.1% SSC, 55 °C) was autoradiographed at -70°C . b, A genomic clone hybridizing to 32 P-pM β 23Cr1 was isolated from a mouse genomic library (partial *Hae*III-digested, *Eco*RI-linked BALB/c embryonic mouse DNA cloned into λ bacteriophage Charon 4A, given by P. Leder) by filter hybridization⁵⁴ and was digested with *Eco*RI, electrophoresed (1 μ g per lane), transferred, hybridized with 32 P-pM β 23Cr1 and autoradiographed as in a. The molecular weight markers are *Hind*III-digested λ DNA fragments.

indicating that the gene for the 23K β -crystallin polypeptide is present in λ M β 23Cr2. The additional DNA bands hybridizing weakly to 32 P-pM β 23Cr1 in the *Eco*RI-digested λ M β 23Cr2 (Fig. 1b) are probably incomplete digestion products containing β 23 gene sequences. The murine DNA fragment inserted into λ M β 23Cr2 is ~ 18 kb long. Restriction mapping indicated that the *Eco*RI 3.2- and 2.3-kb bands in λ M β 23Cr2 are each unique fragments of DNA containing sequences present in pM β 23Cr1.

To determine whether the 3.2- and 2.3-kb *Eco*RI fragments in λ M β 23Cr2 contain different halves of the same gene, they were hybridized with 32 P-probes derived from the 5' or 3' region of the cloned cDNA in pM β 23Cr1. Only the 2.3-kb fragment hybridized to a 96-bp *Pst*I-*Msp*I fragment from the 5' end of the β 23 cDNA and only the 3.2-kb fragment hybridized to a 179-bp *Rsa*I-*Xba*I fragment from the 3' end of the cloned cDNA (data not shown). This indicates that one copy of the β 23 gene was present in the two *Eco*RI genomic DNA fragments.

The 2.3-kb *Eco*RI fragment in λ M β 23Cr2 was subcloned into pBR322 for sequencing. The strategy and partial sequence are shown in Fig. 2. Amino acid residues 3–41 derived from the nucleotide sequence shown in Fig. 2 match the first 39 residues of the 23K β -crystallin polypeptide deduced previously from the cDNA clone¹⁵. ATG (boxed in Fig. 2), present 45 nucleotides upstream from the codon corresponding to the start of the cDNA¹⁵, seems to be the translation initiation codon, as it is the first ATG codon in the sequence followed by an open reading frame continuing into the known coding region for the polypeptide.

The transcription initiation site (cap site) was determined by an S_1 mapping experiment using a 1.3-kb *Eco*RI-*Msp*I gene fragment and polyadenylated mouse lens RNA (Fig. 3). The results show that the mRNA cap site (Figs 2, 3, arrow) is at an A residue 20 nucleotides upstream from the putative initiation site for translation mentioned above. This translation initiation codon is the first ATG sequence downstream from the cap site. The sequence TATGAAA (underlined in Fig. 2), resembling the eukaryotic promoter sequence TATA (Goldberg-Hogness box), is present 30 nucleotides upstream from the cap site. The

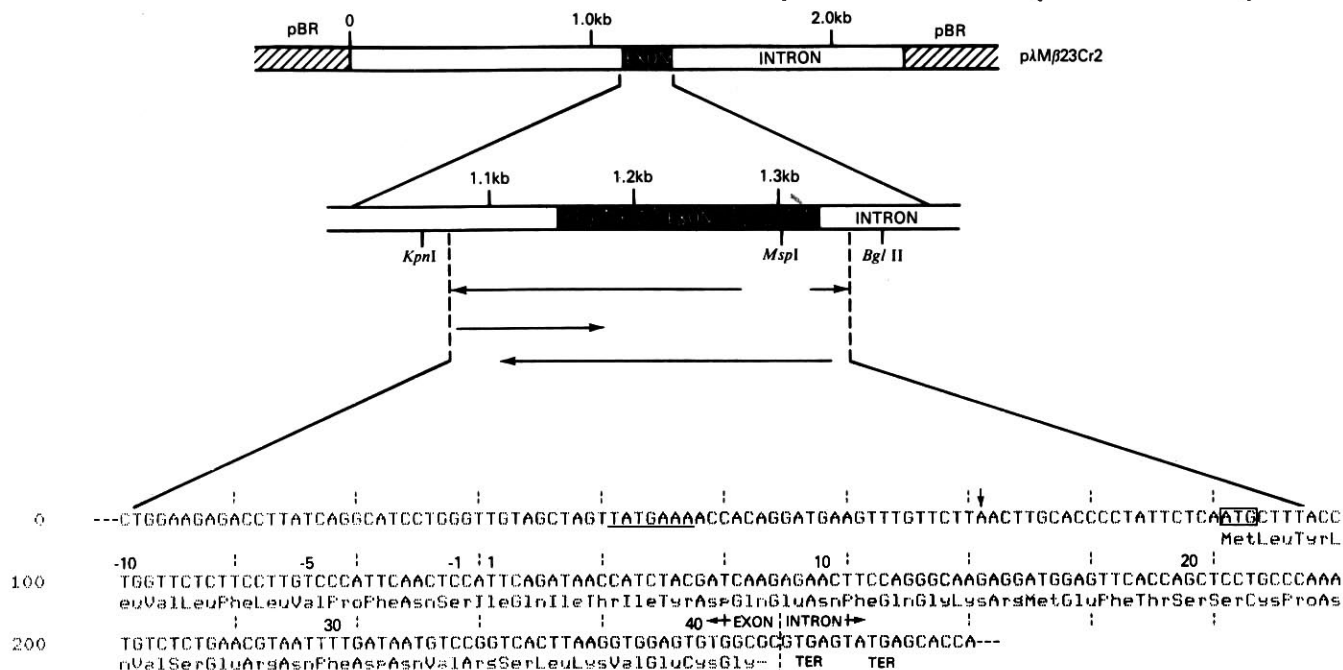
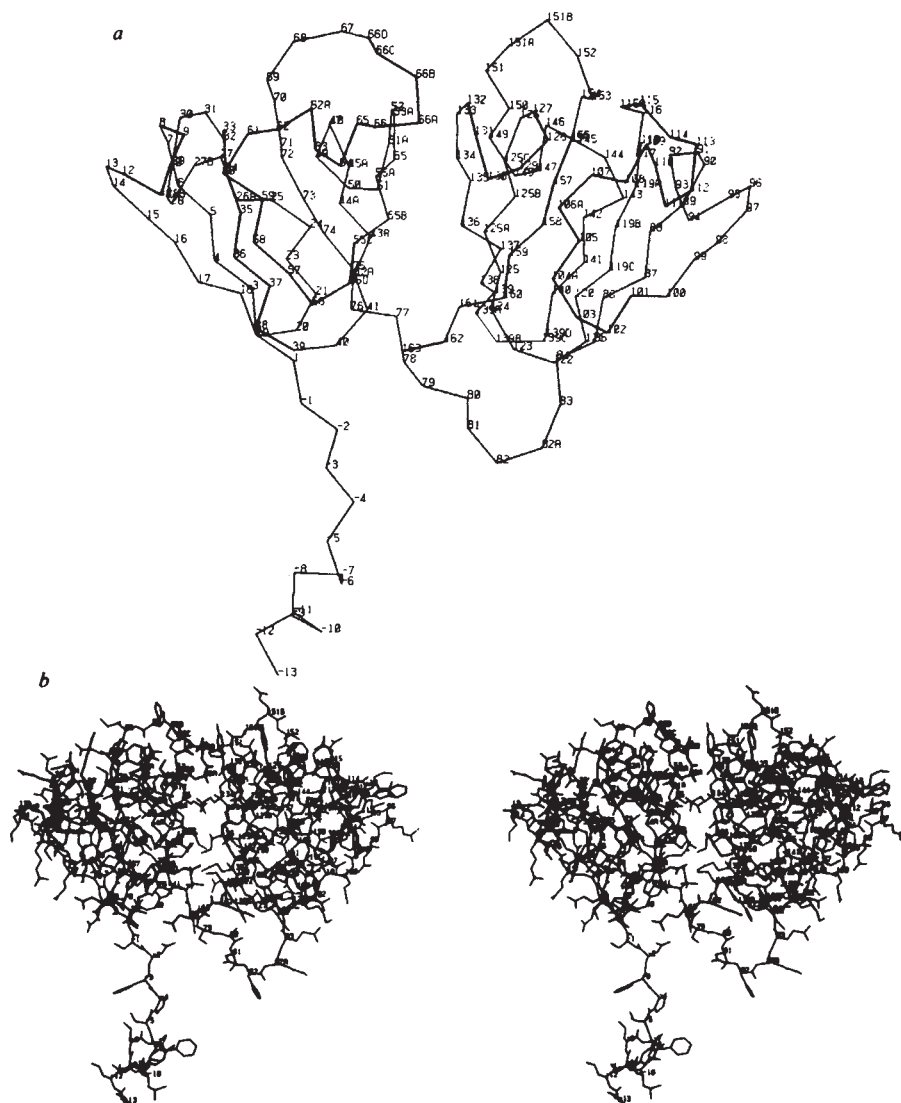


Fig. 2 Nucleotide sequence of the 5' region of the cloned murine 23K β -crystallin gene. A pBR322 subclone (p λ M β 23Cr2) containing the 2.3-kb *Eco*RI fragment of the original genomic clone, λ M β 23Cr2, was analysed by restriction enzyme mapping and DNA sequencing. We used a nick-translation method of DNA sequencing using dideoxynucleotides and DNA restriction fragments labelled at only one of the two 5' ends⁵⁵. Top, schematic diagram of the subclone p λ M β 23Cr2 showing the positions of the first exon and part of the first intron of the 23K β -crystallin gene. Middle, an expanded region around the first exon showing the restriction enzyme sites used for DNA sequencing and the actual regions and directions sequenced. Bottom, the nucleotide sequence and derived amino acid sequence of the region described above; the numbers above the residues designate amino acid positions with position 1 being the Ile equivalent to the first amino acid residue of calf γ II-crystallin; see text for explanation of other designations. TER, termination codon.

Fig. 5 Computer-derived model of the murine 23K β -crystallin polypeptide. *a*, The predicted C α backbone of murine β -crystallin showing the overall arrangement of the polypeptide chain. The numbering system of the murine chain is based on the modified γ_{II} sequence suggested by X-ray analysis¹⁸ with negative numbers referring to the N-terminal extension. *b*, A stereo diagram of the complete molecule. An Evans and Sutherland Picture System II linked to a PDP 11/60 computer was used with the program FRODO⁵⁸, modified for the Picture System by Drs A. Jones and I. J. Tickle, to construct the molecular model and MIDAS (I. J. Tickle, unpublished work) to display the complete molecular structure. We first constructed a main chain having the same conformation as γ_{II} assuming the alignment in Fig. 4 (see also ref. 18). Insertions and deletions were accomplished so as to minimize disruption of the structure (see also ref. 19), then the residues of the hydrophobic core were added and adjusted to occupy positions giving allowed intermolecular distances and optimal van der Waals' interactions. The surface residues were then added in positions close to those occupied by the topologically equivalent residues of γ_{II} , but were occasionally adjusted to optimize interactions, for example to form ion pairs. Finally, the N-terminal residues were added in a conformation suggested by secondary structure prediction techniques.



residues having side chains, while the serines have topologically equivalent positions inaccessible to solvent with the serine O γ hydrogen-bonded to a main-chain NH which stabilizes the Greek key structural motif. Certain other residues are conservatively varied between the four structural motifs, for example the residues equivalent to Tyr 6, Gly 7, Phe 11 and Val 37. The conservative variation of these residues allows the conformation to be close to that of γ -crystallin.

A computer-derived model shown in Fig. 5 indicates several interesting compensatory changes in the murine β -crystallin. For example, the substitution of tryptophan for tyrosine at position 62 is allowed by an insertion at residues 66A, B, C and D which loops out over the tryptophan (see Fig. 5a). There are also some changes in the residues between the domains,

for example the insertion of a serine for a phenylalanine at position 84 which allows the two domains to move closer together.

Of particular interest are the chain termini. Unlike bovine β Bp, the murine β 23 C-terminus stops at the surface of the C-terminal domain. By contrast, there is a putative N-terminal extension (residues -13 to -1) which cannot be integrated into the three-dimensional fold. This N-terminal sequence is encoded in the first exon of the gene and is presumably present in the β -crystallin polypeptide. It remains possible, however, that the N-terminal extension is cleaved from the polypeptide after translation. The N-terminal residues have a very hydrophobic nature with a sequence of Leu, Val, Phe characteristic of membrane anchors and signal peptides which dissolve in the

SV REI	420	425	430	435
	LeuLeuIleIleGlyLeuMetIlePheAlaCysSerMetMetLeuThrSerThrArgArgCOOH			
μ chain IgM	570	575	580	590
	LeuTrpThrThrAlaSerThrPheIleValLeuPheLeuLeuSerPheTyrSerThrThrValThrLeuPheLysValLysCOOH			
Murine β 23K	-13	-12	-11	-10
	MetLeuTyrLeuValLeuPheLeuValProPheAsnSerIleGlnIleThrIleTyr			
Cytochrome P ₄₅₀ LM ₂	1	5	10	15
	MetGluPheSerLeuLeuLeuLeuAlaPheLeuAlaGlyLeuLeuLeuLeuPhe			
Sucrase-isomaltase	10	15	20	25
	IleThrLeuIleValLeuPheValIleValIleIleAlaIleAlaLeuIleAla			
Proinsulin II	-24	-20	-15	-10
	MetAlaLeuTrpIleArgPheLeuProLeuAlaLeuLeuIleLeuTrpGluProArgProAlaGlnAla			
Bacteriorhodopsin	-13	-10	-5	-1
	MetLeuGluLeuLeuProThrAlaValGluGlyValSer			

Fig. 6 Comparison of a hydrophobic region of amino acids at the N-terminus of β 23 with the putative C-terminal transmembrane anchor sequences of the Sindbis virion glycoprotein SVRE1⁵⁹, and the membrane-bound immunoglobulin μ chain, IgM⁶⁰; the putative anchor regions near the N-terminus of sucrase-isomaltase⁶¹ and cytochrome P₄₅₀ LM₂ (ref. 62); and the precursor regions of rat proinsulin II⁶³ and bacteriorhodopsin³³.

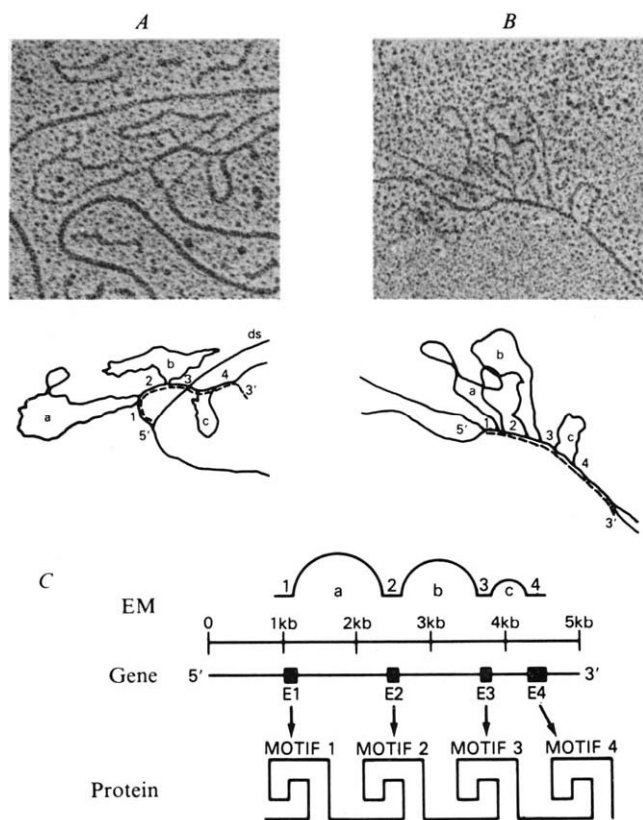


Fig. 7 Electron microscopic analysis of the 23K β -crystallin gene using murine lens polyadenylated RNA (A) and pM β 23Cr1 (B). **A**, Two μ g of the 23K β -crystallin genomic clone DNA (λ M β 23Cr2) and 1 μ g of total murine lens polyadenylated RNA were hybridized at 53 °C overnight in 20 μ l of 70% deionized formamide, 0.1 M tricine pH 7.6, 1 mM EDTA and 0.5 M NaCl in a sealed capillary tube, diluted with the hybridization buffer, prepared for electron microscopy⁶⁴ and examined with a Philips 400 electron microscope. An electron micrograph of a hybrid containing R-loops is shown with an illustrative tracing beneath. The letters designate introns, the numbers exons; the 5' to 3' direction of the mRNA sequence (dotted line) is shown. ds, Displaced DNA strand. **B**, 0.5 μ g of the genomic clone DNA (λ M β 23Cr2) and 1 μ g of the 23K β -crystallin cDNA (pM β 23Cr1) in 100 μ l of 70% deionized formamide, 0.1 M Tris-HCl pH 7.5, 10 mM EDTA and 0.16 M NaCl were heated at 85 °C for 20 min, incubated at 43 °C for 2 h, prepared for electron microscopy and examined as described above. An electron micrograph of a heteroduplex is shown with an illustrative tracing beneath. The letters designate introns, the numbers exons; the 5' to 3' direction of the mRNA sequence encoded in the cDNA (dotted line) is shown. **C**, Schematic diagram showing the electron micrograph (top) and the deduced structure of the 23K β -crystallin gene (middle) with its relationship to the proposed structure of the 23K β -crystallin polypeptide which is divided into four Greek key motifs (bottom). a-c, Introns; 1-4, exons; E1-E4, exons 1 to 4.

membrane hydrophobic lipid fraction. If the N-terminus is buried in a membrane it would probably be largely α -helical²³. Leucines are strongly helix-inducing. However, if the strand were in a β -sheet it would have an interesting arrangement of residues with four leucines on one side at positions -12, -10, -8 and -6 and alternating aromatics and smaller aliphatics on the other side. We have constructed the model with an α -helix between -13 and -5 which is terminated by the proline at -4. If the N-terminus were an anchor, the hydrophilic residues (Ser -1 and Asn -2) would form a flexible link allowing the hydrophobic region access to the lipid bilayer. Homology with a selection of putative hydrophobic membrane anchors is shown in Fig. 6.

Crystallins are close-packed structural proteins. In view of the morphology of eye lenses, although intracellular, some

crystallins might interact with a very extensive membrane surface. The hydrophobic N-terminal peptide of β 23 could link the intracellular bilobal structure to the membrane. Unlike the other anchor membrane sequences in Fig. 6, this murine sequence would be predicted to dip into the membrane without penetrating the extracellular side of the membrane polar head groups. An association of β -crystallin polypeptides with cellular membranes through a hydrophobic extension may contribute to the osmotic balance within the lens. For example, numerous cataracts are associated with both altered ion permeabilities and loss of β -crystallin polypeptides (see ref. 24). One cataract of particular interest in this regard occurs in the Philly mouse²⁵. A 27K β -crystallin polypeptide is absent from the Philly lens^{26,27} in which Na^+ and K^+ permeability is seriously disturbed.

It is also possible that the putative hydrophobic N-terminal peptide of the murine β -crystallin associates with other β -crystallin polypeptides to form polymeric proteins. Hydrophobic bonds are believed to have an important role in the aggregation of the α A-crystallin²⁸⁻³⁰ and δ -crystallin^{31,32} polypeptides. Alternatively, the sequence may be cleaved off as a signal peptide of a pre-protein as mentioned above. This seems unlikely as crystallins are not secreted proteins and the chain is much shorter than most signal peptides (see Fig. 6). One fascinating homology is shown by the N-terminal 13 residues of bacterial prherhodopsin (Fig. 6), a precursor of an integral membrane protein, bacteriorhodopsin. Although these residues are cleaved off in normal growth conditions, the sequence is different from signal sequences and the function is unknown³³.

Domains, structural motifs and exons

The organization of globular proteins in hierarchies of structures—secondary, folding units, supersecondary, domains, tertiary and quaternary—may reflect their evolution^{34,35}. Gilbert³⁶ first suggested that the exons of eukaryotic genes might correspond to polypeptide segments with specific functions. Blake³⁷ has pointed out that if exons also correspond to folded proteins, the new protein is more likely to achieve structural stability by being a combination of folded substructures. This would explain the discovery of different folding units or domains of common structure and function combined with units of different structure and function such as in the dehydrogenases and in triosephosphate isomerase, pyruvate kinase and 2-keto-3-deoxy-6-phosphogluconic aldolase (for review see ref. 35). The gene structures of immunoglobulins^{38,39}, globins⁴⁰⁻⁴⁵ and ovomucoid⁴⁶ serve as examples for the idea that separate exons may encode the structural and functional domains of proteins.

In the β/γ class of crystallins, the protein is organized into four very similar Greek key structural motifs. If these evolved by gene multiplication and fusion it is possible that introns occur between the structural motifs. To test this possibility we examined λ M β 23Cr2 by electron microscopy after hybridization to lens mRNA (Fig. 7A) or to the cloned cDNA (pM β 23Cr1) used to isolate this gene (Fig. 7B). Previous experiments¹⁵, together with the sequencing data shown in Fig. 2, show that pM β 23Cr1 lacks only 65 nucleotides at its 5' end. Three introns were found in the cloned murine β 23 gene, creating four exons. The presence of a poly(A) tail near the smallest intron (c) in the R-loop (Fig. 7A) indicates the 3' end of this gene. Measurements of the four exons (1-4) and three introns (a-c) in 35 R-loops with a numonics digitizer gave the following values: 5'-exon 1 (193 ± 65 bp)—intron a ($1,684 \pm 130$ bp)—exon 2 (147 ± 36 bp)—intron b ($1,433 \pm 157$ bp)—exon 3 (151 ± 29 bp)—intron c (490 ± 44 bp)—exon 4 (184 ± 85 bp)—3'. The size of exon 1 is known by sequencing to be 177 bp (see Fig. 2). The value for the size of exon 4 is less reliable than those for exons 2 and 3 as its 3' border is not delineated by an intron. The predicted sizes of the coding units for motifs 1-4 in our model of β 23 are 159, 159, 126 and 150 bp, respectively, and fall within the measured sizes of their corresponding exons. In addition, the sequence of exon 1 ends exactly at the boundary between the first and second structural motifs of the putative tertiary structure of β 23 (see Figs 2, 4). These data suggest

strongly that each exon encodes a separate structural motif of this murine β -crystallin polypeptide (Fig. 7C).

In addition to linking gene structure to protein structure, the number and positions of introns in this murine β -crystallin gene strengthen the argument that the structure of murine β 23 is homologous to that of bovine γ II. As a computer-derived model of bovine β Bp also resembles that of bovine γ II (ref. 19), it is possible that all the β - and γ -crystallin polypeptides have strongly related tertiary structures exhibiting the four-motif symmetry. It is possible that other β - and γ -crystallin genes have comparably situated exons and introns reflecting the struc-

ture of their respective polypeptides. By contrast, the avian genes for δ -crystallin have 13–15 introns^{47–49}, suggesting that this lens protein has a polypeptide structure very different from that of the β - or γ -crystallins. It is interesting, however, that native δ -crystallin is a tetramer³¹ containing four similar polypeptides^{50,51} which seem to lie on an axis having 2-fold symmetry⁵² in contrast to the monomeric γ -crystallins containing four symmetrically related structural motifs¹⁸.

We thank Dr G. Wistow for critical discussions, Dr P. Leder for the mouse DNA library, and Mrs Florine Carpenter for typing the manuscript.

Received 7 September 1982; accepted 5 January 1983.

1. Harding, J. J. & Dilley, K. J. *Expl Eye Res.* **22**, 1–73 (1976).
2. Bloemendal, H. (ed.) *Molecular and Cellular Biology of the Eye Lens* (Wiley, New York, 1981).
3. Clayton, R. M. in *The Eye Vol. 5* (eds Davson, H. & Graham, L. T.) 399–494 (Academic, London, 1974).
4. Williams, L. A. & Piatigorsky, J. *Eur. J. Biochem.* **100**, 349–357 (1979).
5. Papaconstantinou, J. *Science* **156**, 338–346 (1967).
6. Piatigorsky, J. *Differentiation* **19**, 134–153 (1981).
7. de Jong, W. W. et al. *Eur. J. Biochem.* **53**, 237–242 (1975).
8. de Jong, W. W., Gleaves, J. T. & Boulter, D. *J. molec. Evol.* **10**, 123–135 (1977).
9. Wilson, A. C., Carlson, S. S. & White, T. J. *A. Rev. Biochem.* **46**, 573–639 (1977).
10. Piatigorsky, J. *Biochemistry* **20**, 6427–6431 (1981).
11. de Jong, W. W., Stapel, S. O. & Zweers, A. *Comp. Biochem. Physiol.* **69B**, 593–598 (1981).
12. Yu, N. T., East, E. J. & Chang, R. C. *Expl Eye Res.* **24**, 321–334 (1977).
13. Horwitz, J. & Piatigorsky, J. *Biochim. biophys. Acta* **624**, 21–29 (1980).
14. Driessen, H. P. C., Herbrink, P., Bloemendal, H. & de Jong, W. W. *Eur. J. Biochem.* **121**, 83–91 (1981).
15. Inana, G., Shinohara, T., Maizel, J. V. Jr & Piatigorsky, J. *J. biol. Chem.* **257**, 9064–9071 (1982).
16. Croft, L. R. *Biochem. J.* **128**, 961–970 (1972).
17. Dayhoff, M. O. *Atlas of Protein Sequence and Structure Vol. 5*, Suppl. 3 (National Biomedical Research Foundation, Silver Spring, Maryland, 1978).
18. Blundell, T. L. et al. *Nature* **289**, 771–777 (1981).
19. Wistow, G. et al. *FEBS Lett.* **133**, 9–16 (1981).
20. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
21. Hejtmancik, J. F. *J. biol. Chem.* (in the press).
22. Breathnach, R. & Chambon, P. *A. Rev. Biochem.* **50**, 349–383 (1981).
23. Engelman, D. M. & Steitz, T. A. *Cell* **23**, 411–422 (1981).
24. Piatigorsky, J. *Curr. Topics Eye Res.* **3**, 1–39 (1980).
25. Kador, P. F., Fukui, H. N., Fukushi, S., Jernigan, H. M. & Kinoshita, J. H. *Expl Eye Res.* **30**, 59–68 (1980).
26. Zigler, J. S., Carper, D. A. & Kinoshita, J. H. *Ophthalm. Res.* **13**, 237–251 (1981).
27. Carper, D., Shinohara, T., Piatigorsky, J. & Kinoshita, J. H. *Science* **217**, 463–464 (1982).
28. Li, L.-K. & Spector, A. *Expl Eye Res.* **13**, 110–119 (1972).
29. Van der Ouderdaa, F. J., de Jong, W. & Bloemendal, H. *Eur. J. Biochem.* **39**, 207–222 (1973).
30. de Jong, W. W. et al. *Eur. J. Biochem.* **53**, 237–242 (1975).
31. Piatigorsky, J., Zelenka, P. & Simpson, R. T. *Expl Eye Res.* **18**, 435–446 (1974).
32. Piatigorsky, J., Horwitz, J. & Simpson, R. T. *Biochim. biophys. Acta* **490**, 279–289 (1977).
33. Chang, S. H. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3398–4302 (1981).
34. Rossmann, M. G., Moras, D. & Olsen, K. W. *Nature* **250**, 194–199 (1974).
35. Schulz, G. E. & Schirmer, R. H. *Principles of Protein Structure* (Springer, Heidelberg, 1978).
36. Gilbert, W. *Nature* **271**, 501 (1978); *ICN-UCLA Symp. molec. cell. Biol.* **14**, 1–12 (1979).
37. Blake, C. C. F. *Nature* **273**, 267 (1978).
38. Toneygawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1485–1489 (1978).
39. Sakano, H. et al. *Nature* **277**, 627–633 (1979).
40. Tilghman, S. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **75**, 725–729 (1978).
41. Jeffreys, A. J. & Flavell, R. A. *Cell* **12**, 1097–1108 (1977).
42. Leder, A. et al. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6187–6191 (1978).
43. Craik, C. S., Buchman, S. R. & Beychok, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1384–1388 (1980).
44. Eaton, W. A. *Nature* **284**, 183–185 (1980).
45. Go, M. *Nature* **291**, 90–92 (1981).
46. Stein, J. P., Catterall, J. F., Kristo, P., Means, A. R. & O'Malley, B. W. *Cell* **21**, 681–687 (1980).
47. Bhat, S. P., Jones, R. E., Sullivan, M. A. & Piatigorsky, J. *Nature* **284**, 234–238 (1980).
48. Jones, R. E., Bhat, S. P., Sullivan, M. A. & Piatigorsky, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5879–5883 (1980).
49. Yasuda, K., Kondoh, H., Okada, T. S., Nakajima, N. & Shimura, Y. *Nucleic Acids Res.* **10**, 2879–2891 (1982).
50. Piatigorsky, J. *J. biol. Chem.* **251**, 4416–4420 (1976).
51. Reszelbach, R., Shinohara, T. & Piatigorsky, J. *Expl Eye Res.* **25**, 583–593 (1977).
52. Narebor, E., Slingsby, C., Lindley, P. F. & Blundell, T. L. *J. molec. Biol.* **143**, 223–225 (1980).
53. Maniatis, T., Jeffrey, A. & Kleid, D. C. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).
54. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961–3965 (1975).
55. Seif, I., Khoury, G. & Dhar, R. *Nucleic Acids Res.* **8**, 2225–2240 (1980).
56. Berk, A. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1274–1278 (1978).
57. Wistow, G. thesis, Univ. London (1982).
58. Jones, T. A. *J. appl. Crystallogr.* **11**, 268–272 (1978).
59. Rice, C. M., Bell, J. R., Hunkapiller, M. W., Strauss, E. G. & Strauss, J. H. *J. molec. Biol.* **154**, 355–378 (1982).
60. Rogers, J. et al. *Cell* **20**, 303–312 (1980).
61. Frank, G. et al. *FEBS Lett.* **96**, 183–188 (1978).
62. Haugen, D. A., Armes, L. G., Yasunobu, K. T. & Coon, M. J. *Biochem. biophys. Res. Commun.* **77**, 967–973 (1977).
63. Lomedico, P. et al. *Cell* **18**, 545–558 (1979).
64. Bhat, S. P. & Piatigorsky, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3299–3303 (1979).

LETTERS TO NATURE

A test for transverse motions of clusters of galaxies

M. Birkinshaw & S. F. Gull

Mullard Radio Astronomy Observatory, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

Although a static lens can produce no change in the brightness of an isotropic radiation field, a moving lens produces a characteristic two-sided brightness pattern that may be used to deduce the transverse velocity of the lens. The magnitude of the effect is proportional to both the transverse velocity and the strength of the lens. We suggest here that this effect provides a method for measuring the transverse velocities of clusters of galaxies; in this case the lens is gravitational and the radiation field is the isotropic microwave background. The predicted magnitude of this effect is at the present sensitivity limits of observation but is at least as large as the expected level of primordial background fluctuations.

The physical phenomenon involved can easily be understood: consider the appearance of an isotropic background radiation viewed through a convex lens moving with velocity v perpendicular to the line of sight (Fig. 1a). Then photons coming from

the side of the lens towards which the lens is moving have had their energies slightly decreased by deflection at the moving lens, the energy change being approximately proportional to $v/c \times$ (deflection angle). The deflected photons therefore have energies appropriate to a different brightness from that of undeflected photons. In the Rayleigh–Jeans part of a black-body spectrum the brightness increases with frequency, so the deflected photons appear at a higher brightness than undeflected photons of the same observed frequency. Similarly, on the opposite side of the lens, the deflected photons gain energy, so that a brightness change of the opposite sign is produced. This characteristic two-sided brightness structure (Fig. 2) can be used to measure the magnitude and direction of the transverse velocity component of the lens. We now make a more precise formulation of this effect and estimate its magnitude for the case of a cluster of galaxies.

Suppose that an observer, at rest at the origin of coordinates, views a lens with peculiar velocity

$$v = \beta c (\sin \alpha, 0, \cos \alpha) \quad (1)$$

and consider the history of a photon (segment of wavefront) from a background radiation (isotropic in the frame of the observer) with frequency ν and wave vector

$$\mathbf{k} = -\frac{2\pi\nu}{c} (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \quad (2)$$

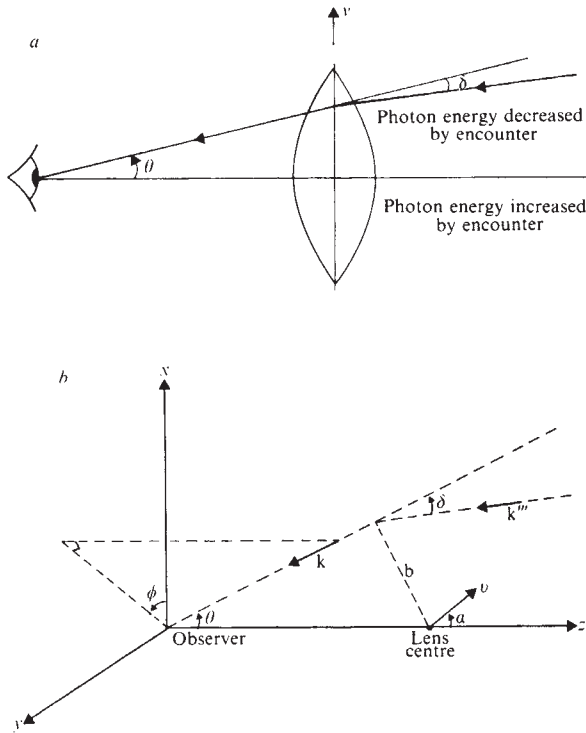


Fig. 1 *a*, The deflection of light by a moving lens leads to a slight decrease in the photon energy in $\theta > 0$, and a slight increase in $\theta < 0$. In the Rayleigh-Jeans part of the spectrum of a background radiation, this appears as a slight relative brightness increase in $\theta > 0$, and an decrease in $\theta < 0$. *b*, The apparent configuration of the problem in the frame of the observer. The lens has velocity $\mathbf{v}$ in the x - z plane, and the angles (θ, ϕ) describing the direction of the observed (deflected) photon correspond to the angle between $\mathbf{k}$ and the z -axis and the projected $\mathbf{k}$ and the projected $\mathbf{v}$, respectively.

which arrives at the observer at the instant that the lens centre is located at

$$\mathbf{x}_L = (\beta d \sin \alpha, 0, d(1 + \beta \cos \alpha)) \quad (3)$$

so that θ is the angle between $\mathbf{k}$ and $\mathbf{x}_L$, measured by the observer and ϕ is the angle between the $\mathbf{k}$ -vector and $\mathbf{v}$ -vector projected on the plane of the sky (Fig. 1*b*). Then the photon appears in the frame of the lens (the 'primed' frame) with frequency ν' and wave-vector $\mathbf{k}'$ where $\gamma = (1 - \beta^2)^{-1/2}$,

$$\begin{aligned} \nu' &= \gamma \nu (1 + \beta \theta \sin \alpha \cos \phi + \beta \cos \alpha) \\ \mathbf{k}' &= -\frac{2\pi\nu}{c} [\gamma \beta \sin \alpha + (\gamma - 1) \sin \alpha \cos \alpha \\ &\quad + \theta \cos \phi (\gamma \sin^2 \alpha + \cos^2 \alpha)], \\ &\quad \theta \sin \alpha, \\ &\quad [\gamma \beta \cos \alpha + (\gamma \cos^2 \alpha + \sin^2 \alpha) \\ &\quad + \theta (\gamma - 1) \sin \alpha \cos \alpha \cos \phi] \end{aligned} \quad (4)$$

to first order in the small angle θ . At the instant of arrival of the photon at the observer, the lens is at

$$\mathbf{x}'_L = d[(\gamma - 1) \sin \alpha \cos \alpha + \gamma \beta \sin \alpha], 0, [\gamma \cos^2 \alpha + \sin^2 \alpha + \gamma \beta \cos \alpha] \quad (5)$$

in its rest frame. The impact parameter for the photon-lens encounter is

$$\mathbf{b}' = \mathbf{k}' \left(\frac{\mathbf{k}' \cdot \mathbf{x}'_L}{\mathbf{k}' \cdot \mathbf{k}'} \right) - \mathbf{x}'_L \quad (6)$$

in this frame. The deflection at this impact parameter, $\delta'(b') = \delta'(d\theta)$ to order θ , is given by the lens equation. The effect of

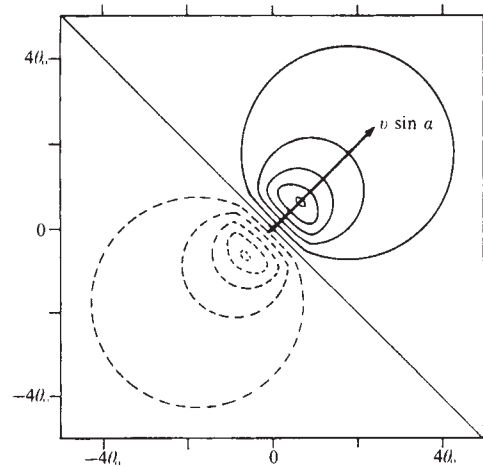


Fig. 2 Contours of ΔT_{RJ} for the lens model described in the text. The units of θ are $\theta_0 = (R_0/d)$, and contours are drawn at intervals of $0.18 \Delta T_{RJ}^{\max}$ from $-0.9 \Delta T_{RJ}^{\max}$ to $+0.9 \Delta T_{RJ}^{\max}$ with negative contours drawn broken.

the encounter is to rotate the $\mathbf{k}''$ vector to $\mathbf{k}'$, where the wave-vector before encounter in the frame of the lens, $\mathbf{k}''$, is

$$\mathbf{k}'' = \mathbf{k}' + \delta'(b') \mathbf{b}' \left(\frac{\mathbf{k}' \cdot \mathbf{k}'}{\mathbf{b}' \cdot \mathbf{b}'} \right)^{1/2} \quad (7)$$

to order δ' . The photon frequency is unchanged in this frame, but in the frame of the observer the frequency of the photon before encounter is

$$\nu'' = \gamma \left(\nu' + \frac{\mathbf{v} \cdot \mathbf{k}''}{2\pi} \right) \quad (8)$$

so that to order δ' and θ , the fractional change in frequency caused by the lensing effect is

$$\frac{\Delta \nu}{\nu} = \frac{\nu'' - \nu}{\nu} = \beta \gamma \sin \alpha \cos \phi \delta'(b') \quad (9)$$

The corresponding change of brightness temperature of the radiation field is

$$\Delta T_{RJ} = \frac{c^2}{2k\nu} \left(\frac{\Delta \nu}{\nu} \right) \left(\frac{\partial B_\nu}{\partial \nu} \right) \quad (10)$$

where B_ν is the brightness of the background radiation.

It is of interest to estimate the size of the effect predicted by equations (9) and (10) for a gravitational lens moving across the microwave background radiation. We consider a transparent gravitational lens with mass M_0 distributed in a constant-density spheroid of radius R_0 . Then, to order $(GM_0/c^2 R_0)$, the field equations of general relativity can be solved to give

$$\delta'(b') = \frac{4GM_0}{c^2 R_0} \begin{cases} \frac{\theta_0}{\theta} \left[1 - \left(1 - \frac{\theta^2}{\theta_0^2} \right)^{3/2} \right] & \theta \leq \theta_0 \\ \frac{\theta_0}{\theta} & \theta \geq \theta_0 \end{cases} \quad (11)$$

where $\theta_0 = R_0/d$, so that, for observations in the Rayleigh-Jeans part of the microwave background spectrum,

$$\begin{aligned} \Delta T_{RJ}(\theta, \phi) &= \left(\frac{8\gamma \beta G M_0 T_r \sin \alpha}{c^2 R_0} \right) \\ &\quad \times \cos \phi \begin{cases} \frac{\theta_0}{\theta} \left[1 - \left(1 - \frac{\theta^2}{\theta_0^2} \right)^{3/2} \right] & \theta \leq \theta_0 \\ \frac{\theta_0}{\theta} & \theta \geq \theta_0 \end{cases} \end{aligned} \quad (12)$$

The resulting angular dependence $\Delta T_{RJ}(\theta, \phi)$ is illustrated in Fig. 2. The maximum $|\Delta T_{RJ}|$, $\Delta T_{RJ}^{\max}$, occurs at $\theta = \theta_0$ and $\phi = 0$

or π , and is

$$(\Delta T_{RJ}^{\max}/\text{mK}) = 0.04(v/1,000 \text{ km s}^{-1})(M_0/10^{15} M_\odot) \times (R_0/100 \text{ kpc})^{-1} \sin \alpha \quad (13)$$

for $\beta \ll 1$. Since the present limits of sensitivity are typically 0.3 mK (ref. 1), it is clear that the effect will be undetectable with current instrumentation for most lenses. However, more sensitive receivers are coming into use and can hope to achieve sensitivities of <0.1 mK, so that for a rapidly moving cluster of galaxies with $M_0 \sim 3 \times 10^{15} M_\odot$, $R_0 \sim 500 \text{ kpc}$ and $v \sim 6,000 \text{ km s}^{-1}$ for which $\Delta T_{RJ}^{\max} \sim 0.15 \text{ mK}$, the effect may be just detectable in the near future. This effect is of the same order, or larger than, the primordial small angular scale anisotropies of the microwave background radiation².

This simple argument has ignored the effects of the cosmology within which the problem is embedded; extensions to other cosmologies than the flat spacetime assumed here will only modify the result of equation (12) slightly for cluster redshifts ≤ 0.5 .

We thank Bill Shuter for provoking these calculations through his comments at a seminar given by S.F.G. M.B. thanks Gonville and Caius College, Cambridge for a research Fellowship.

Received 6 December 1982; accepted 20 January 1983.

1. Birkinshaw, M., Gull, S. F. & Northover, K. J. E. *Mon. Not. R. astr. Soc.* **197**, 571–592 (1981).
2. Boynton, P. E. *Proc. IAU Symp.* No. 79, 317–324 (1978).

Sub-second pulsations simultaneously observed at microwaves and hard X rays in a solar burst

T. Takakura*, P. Kaufmann†, J. E. R. Costa†, S. S. Degaonkar*‡, K. Ohki§ & N. Nitta*

* Department of Astronomy, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

† Instituto de Pesquisas Espaciais, C.P. 515, 12200, São José dos Campos, SP, Brazil

§ Tokyo Astronomical Observatory, University of Tokyo, Mitaka, Tokyo 181, Japan

Sub-second time structures have been found in millimetre waves^{1,2} and, independently, in hard X rays^{3,4} emitted during solar bursts. However, because simultaneous observations of such fast time structures in millimetre radio and X-ray ranges had not been available, we planned coordinated observations of the solar burst in November 1981 with a high time resolution of a few milliseconds. The hard X rays (30–40 keV) were observed by the hard X-ray monitor, HXM⁵, aboard the Hinotori Satellite with a time resolution of 7.81 ms and the radio emissions were observed on the ground by the 45-ft dish at Itapetinga Radio Observatory with a high time resolution (1 ms) and high sensitivities at 22 GHz and 44 GHz (ref. 6), supplemented by patrol observations at 7 GHz with a time resolution of 100 ms. Absolute timing at Itapetinga and Hinotori is better than 10 ms. Correlated sub-second time structures at hard X-ray and at millimetre wavelengths were found for the first time.

A solar sub-flare occurred on 4 November 1981 at 18 h 28 min 20 s UT. The associated radio and X-ray bursts, shown in compressed time scale in Fig. 1a, b, reveal that the correlation of the major burst time structures is poor among the three microwave frequencies, and between microwaves and hard X

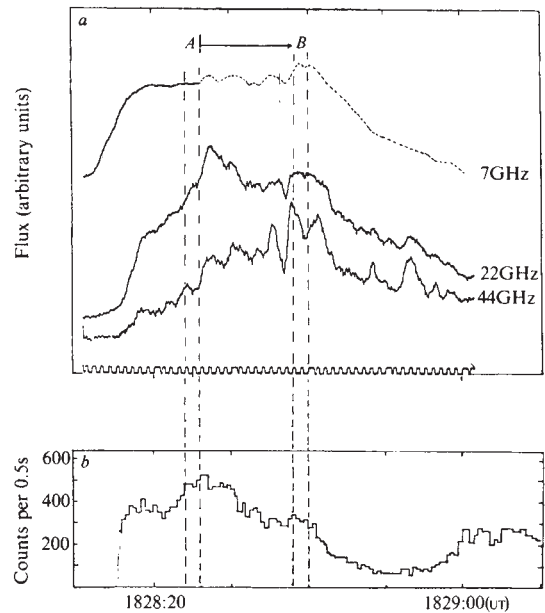


Fig. 1 Time profiles in microwaves (a) and in hard X rays (b) for the solar burst of 4 November 1981. Vertical lines show intervals labelled A and B which are expanded in Figs 2 and 3, showing the superimposed 3–4 s⁻¹ quasi-pulsations.

rays. However, during this solar burst, it was found in time expanded sections that quasi-periodic pulsations were present, repeating at a rate of 3–4 s⁻¹. The pulsations persisted throughout the maximum phase of the burst, that is from about 1828:10–1828:40 UT, at 22 GHz, 44 GHz and at hard X rays. For analysis with greater time resolution we expanded various sections in that time interval. For examples, we show two sections labelled A and B in Fig. 1, each of 2 s duration. Figure 2 shows the gradual component has not been subtracted from curves b and d for both sections A and B. In Fig. 2c and e, the radio flux was filtered by taking the second time derivative to eliminate the contamination due to changes in the gradual background level. The X-ray data were not filtered in this way as the background level was nearly constant in those periods. The statistical error σ of the X-ray counts, given by the square root of an average counting rate, shown in Fig. 2a is large and corresponds to $\sim 50\%$ of the peak-to-peak amplitude of the pulsations. The error bars of the radio signal are due to system temperature and are small as indicated by σ in Fig. 2b and c.

There are six distinctive fast time structures in Fig. 2 for sections A and B and they exhibit a striking correspondence with each other. For most of the time structures, there is nearly one-to-one correspondence. The peak-to-peak amplitude of the X-ray pulsations is $\sim 30\%$ of the average counts, and is larger than the statistical error. On the other hand, the peak-to-peak amplitude of the radio pulsations is $\sim 1\%$ of the gradual background level, and is larger than the system noise error bars. The best signal-to-noise ratio is obtained at 22 GHz.

To test the significance of the correlation, a cross-correlation analysis was made. The cross-correlation coefficients are given by

$$c_j = \frac{1}{n-j+1} \sum_{i=1}^{n-j+1} (x_i - \bar{x})(y_{i+j-1} - \bar{y}) \quad (1)$$

where $\bar{x}$, $\bar{y}$ are the mean values for the time interval under investigation. The errors for the coefficients are derived from the general theory of propagation of errors⁷,

$$\sigma_{c_j}^2 = \sum_k \left[\left(\frac{\partial c_j}{\partial x_k} \right)^2 \sigma_{x_k}^2 + \left(\frac{\partial c_j}{\partial y_k} \right)^2 \sigma_{y_k}^2 \right] \quad (2)$$

‡ Permanent address: Physical Research Laboratory, Ahmedabad, India.

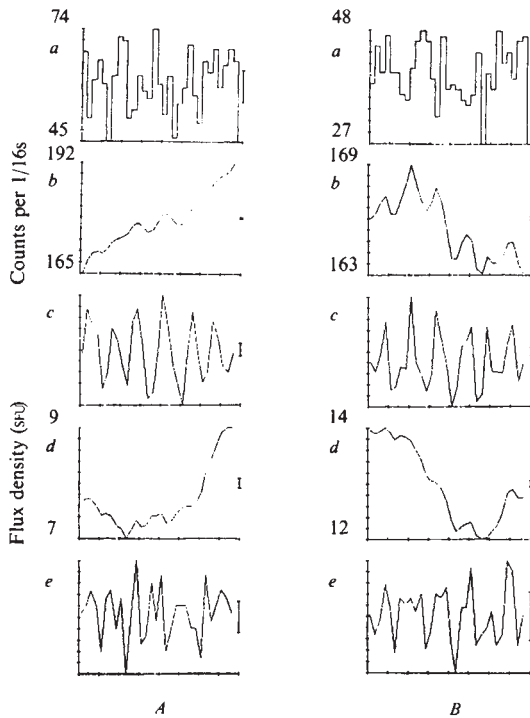


Fig. 2 Expanded time profiles for 2 s in intervals *A* and *B* with an accuracy of 62.5 ms. Displayed plots are for 1.875 s. *a*, X rays; *b* and *c*, 22 GHz; *d* and *e*, 44 GHz.

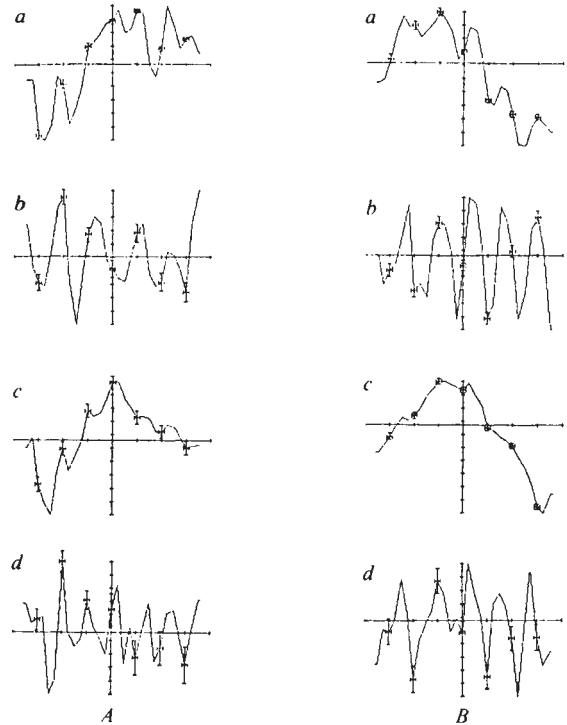


Fig. 3 Normalized cross-correlation coefficients between X rays and microwaves in intervals *A* and *B* for time delays extending to ± 1 s with an accuracy of 62.5 ms. *a* and *b*, X rays, 22 GHz; *c* and *d*, X rays, 44 GHz.

which reduces from equation (1) to

$$\sigma_{c_i}^2 = \frac{1}{(n-j+1)^2} \left\{ \sum_{k=1}^{n-j+1} (y_{k+j-1} - \bar{y})^2 \sigma_{x_k}^2 + (x_{k-j+1} - \bar{x})^2 \sigma_{y_k}^2 \right\} \quad (3)$$

The cross-correlations between X rays and radio fluxes are shown in Fig. 3 for sections *A* and *B*. The abscissa indicates delay time (in seconds) of the radio emission from the X rays. Figure 3*a* and *c* show the cross-correlations for the raw radio data without filtering and Fig. 3*b* and *d*, those with filtered radio data. Clearly the influence of the gradual background at 22 and 44 GHz is significant. Figure 4*A* and *B*, show the cross-correlation between filtered 22 and 44 GHz data for sections *A* and *B*. The accuracy in time delays for these plots of cross-correlations is 62.5 ms.

The quasi-periodicity is evident as the repetition of cross-correlation coefficient peaks occurs in all the plots of Figs 3 and 4. The cross-correlation coefficient plot for 22 and 44 GHz data (Fig. 4) is the best for quasiperiodicity determinations. It gives a period of ~ 290 ms for section *A*, and of ~ 320 ms for section *B*. Periods of the order of 300 ms are also evident on the plots of cross-correlation coefficients in Fig. 3.

The 22 and 44 GHz pulsations appear to be essentially in phase (to within 62.5 ms) for both sections *A* and *B* (Fig. 4). The hard X-ray pulsations, however, do not seem to be coincident in time with the millimetre wave pulsations. Figure 3*a* indicates that 22 GHz peaks might be delayed with respect to the hard X-ray peaks, by about 200 ms and the 44 GHz peaks seemed to be delayed by about 250 ms with respect to hard X-ray peaks, or in phase, or advanced by a small amount (within the time inaccuracy of 62.5 ms). Similarly, in Fig. 3*B*, the millimetre-wave peaks, both at 22 and 44 GHz, might be delayed by 250 ms with respect to hard X rays, or in phase, or advanced by a small amount (within the time inaccuracy of 62.5 ms). It is more likely that the millimetre wave peaks are delayed by an amount close to one period (that is, ~ 300 ms).

Thus, the foregoing observations reveal an interesting correlation of quasi-periodic sub-second pulsations in millimetre

waves and hard X rays. The physical implication of the good correlation with a suggested phase shift of the order of one-period is not clear at present, but it is definitely linked with the understanding of the high-energy phenomena occurring during the solar flares.

If we suppose that the pulsations are attributed to the magnetohydrodynamic (MHD) waves moving with a speed of the order of $10^{3.5}$ km s $^{-1}$, the 300 ms period gives the size of $\sim 10^3$ km for the source showing modulation in intensity. On the other hand, the X-ray image of the flaring site obtained with the hard X-ray telescope SXT 8,9 aboard the Hinotori indicate a single source 10 of about 20 arc s (about 1.5×10^4 km) in HWHM during section *A* of Fig. 1. These measurements have spatial and temporal resolutions of about 10 arc s and 7 s, respectively. The source may, however, be composed of many unresolved small sources as has been proposed by some models 11,12 . Several situations might be conceived. If only one of the smaller sources suffers a modulation in intensity, there must be less than three elements instantaneously present to satisfy such a large modulation as 30%. On the other hand, quasi-periodic reconnections due to tearing mode instability in a magnetic tube has been suggested 13 . The multiple sources, however, need not be interdependent or sequential. They could be attributed to an aggregation of many disrupting flux tubes superimposed on a gradual longer disruption, as has been recently suggested 12 . The superposition would sometimes result in the apparent quasiperiodic fluctuations, whose amplitude depends on the pulse waveform in response to each individual disruption and on how they pile up.

An alternative could be conceived by attributing the period of 300 ms to the bouncing motion of a cloud of high-energy electrons in a magnetic flux tube of the order of 10^4 km in length. A recent model assuming MHD oscillations at the flaring site produces pulsations of microwave and X-ray emission 14 . In this model fluxes of energetic electrons are modulated at the top of a loop, producing hard X-rays at every bounce by bremsstrahlung lower in the chromosphere and modulation of

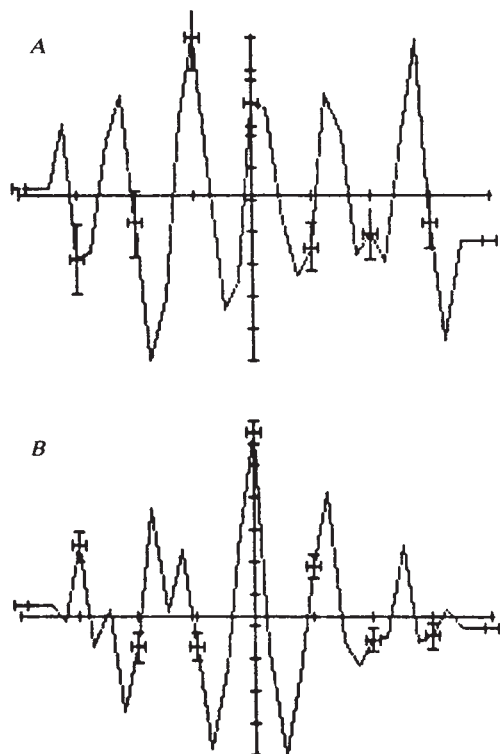


Fig. 4 Normalized cross-correlation coefficients between 22 and 44 GHz in intervals A and B for time delays extending to ± 1 s with an accuracy of 62.5 ms.

microwaves is attributed to the growth rate of Bernstein mode. For producing modulation at such high frequencies as 22 and 44 GHz, however, this model requires very high plasma density of the order of 10^{12} – 10^{13} cm $^{-3}$ at the top of the coronal magnetic tube.

Note that fast microwave time structures superimposed on the event of 4 November 1981 as reported here were also observed at 10.6 GHz at Algonquin Radio Observatory, Canada (K. F. Tapping, personal communication). These data will be included in a more extended study of this event.

We acknowledge the contribution of the late Dr F. M. Strauss (1942–81) to the developments at Itapetinga which made the present high sensitivity observations possible. T.T., K.O. and N.N. thank the staff of the Institute of Space and Astronautical Science, especially to the members of Hinotori team for attention at every stage, from design to launch and operation of the Hinotori Satellite. S.S.D. thanks Professor T. Takakura and JSPS for hospitality. We also thank F. N. S. Carvalho for help in Itapetinga data processing.

Received 13 December 1982; accepted 21 January 1983.

- Kaufmann, P., Strauss, F. M., Opher, R. & Laporte, C. *Astr. Astrophys.* **87**, 58–62 (1980).
- Hurford, G. J., Marsh, K. A., Zirin, H., Kaufmann, P. & Strauss, F. M. *Abstr. Bull. Am. astr. Soc.* **11**, 678 (1979).
- Orwig, L. E., Frost, K. J. & Dennis, B. R. *Astrophys. J. Lett.* **244**, L163–L166 (1981).
- Charikov, Y. E., Kocharov, G. E. & Lazutkov, V. P. *17th int. Cosmic Ray Conf. Paris* (1981).
- Ohki, K. *et al. Proc. Hinotori Symp. on Solar Flares* 69–88 (ISAS, 1982).
- Kaufmann, P., Strauss, F. M., Schaal, R. E. & Laporte, C. *Solar Phys.* **78**, 389–399 (1982).
- Bevington, P. R. *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, New York, 1969).
- Makishima, K. *Proc. Hinotori Symp. on Solar Flares* 120–129 (ISAS, 1982).
- Takakura, T. *et al. Astrophys. J.* (submitted).
- Takakura, T., Ohki, K., Tsuneta, S. & Nitta, N. *Proc. of Japan-US Seminar on Recent Advances in the Understanding of Solar Flares* (1982).
- Brown, J. C., Craig, J. D. & Karpen, J. T. *Solar Phys.* **67**, 143–162 (1980).
- Sturrock, P. A., Kaufmann, P. & Smith, D. F. *Energy Release in Solar Flares* (Stanford University Institute of Plasma Research, Rep. SUIPR No. 933, 1982).
- Spicer, D. S. *Solar Phys.* **70**, 149–172 (1981).
- Zaitsev, V. V. & Stepanov, A. V. *Solar Phys.* (submitted).

Daily observations of a large period jump of the Vela pulsar

P. M. McCulloch*, P. A. Hamilton*, G. W. R. Royle* & R. N. Manchester†

* Physics Department, University of Tasmania, GPO Box 252C, Hobart, Tasmania, Australia 7001

† CSIRO Division of Radiophysics, Epping, New South Wales 2121, Australia

Pulse arrival-time data for the Vela pulsar, PSR0833–45, have been recorded at Hobart for 5 h per day from 8 October 1981. The pulsar has also been monitored at approximately weekly intervals in a continuing programme using the NASA Deep Space Station at Tidbinbilla. The pulse period decreased by about 102 ns between the end of observations at 10 October 1981 2350 UT and the start of observations at 1915 UT the next day¹. We report here that the recovery from this period jump, which has been monitored almost daily at Hobart over a 6-week period, is accurately described by a model consisting of the sum of two exponential decays with time constants respectively 1.62 ± 0.18 days and 233 ± 1 days. The slower decay is consistent with the two-component neutron star model² and has similar parameters to previously observed jumps for this pulsar. However, the rapid initial decay may indicate the presence of a third component which is tightly coupled to the neutron star crust.

Observations at Tidbinbilla were made with a 26-m diameter antenna at a frequency of 2,295 MHz (ref. 3). The pulsar was observed for 22 min, and integrated pulse profiles combined to give a mean pulse arrival time for the day. Uncertainty in mean pulse arrival times is ~ 10 μ s. The Hobart observations are made at a frequency of 645 MHz using a 14-m diameter antenna which can track the source for 5 h. A dual-channel receiver with field-effect transistor first stages is used to give two linear polarizations which are summed after detection; the system noise temperature for each channel is 100 K. Integrated pulse profiles of 1,000 pulses are recorded at intervals of about 2 min throughout the time that the source is visible. The signal-to-noise ratio of each profile is about 20:1, allowing a mean pulse arrival time to be determined with an uncertainty of about 80 μ s. All the arrival time data are subsequently reduced to arrival times at the Solar System barycentre using standard

Table 1 Pre- and post-jump parameters for the Vela pulsar

Pre-jump parameters ($T_0 = \text{JD } 2444795.8399$)	
$\nu = 11.2036049179 \pm 4$ Hz	$P = 0.089256985348 \pm 3$ s
$\dot{\nu} = (-1.55878 \pm 0.00002) \times 10^{-11}$ Hz s $^{-1}$	$\dot{P} = (124.1850 \pm 0.0016) \times 10^{-15}$ s s $^{-1}$
$\ddot{\nu} = (3.9 \pm 0.5) \times 10^{-22}$ Hz s $^{-2}$	$\ddot{P} = (-31 \pm 4) \times 10^{-25}$ s s $^{-2}$
r.m.s. residual = 7 μ s	

Parameters of the jump ($T_j = \text{JD } 2444888.5707 \pm 0.0002$)	
$\Delta\nu_c = (1.0476 \pm 0.0003) \times 10^{-5}$ Hz	
$\Delta\nu_1 = (9.2 \pm 0.9) \times 10^{-8}$ Hz	
$\tau_1 = 1.62 \pm 0.18$ days	
$\Delta\nu_2 = (2.260 \pm 0.003) \times 10^{-6}$ Hz	
$\tau_2 = 233 \pm 1$ days	
r.m.s. residual = 35 μ s	

Post-jump decay parameters		
$\Delta\nu/\nu$	$(8.2 \pm 0.8) \times 10^{-9}$	Long-term decay $(1.137 \pm 0.003) \times 10^{-6}$
$\Delta\dot{\nu}/\dot{\nu}$	$(42 \pm 4) \times 10^{-3}$	$(7.21 \pm 0.06) \times 10^{-3}$
Q	1.0 (assumed)	0.177 $\pm$ 0.001
τ (days)	1.62 ± 0.18	233 $\pm$ 1

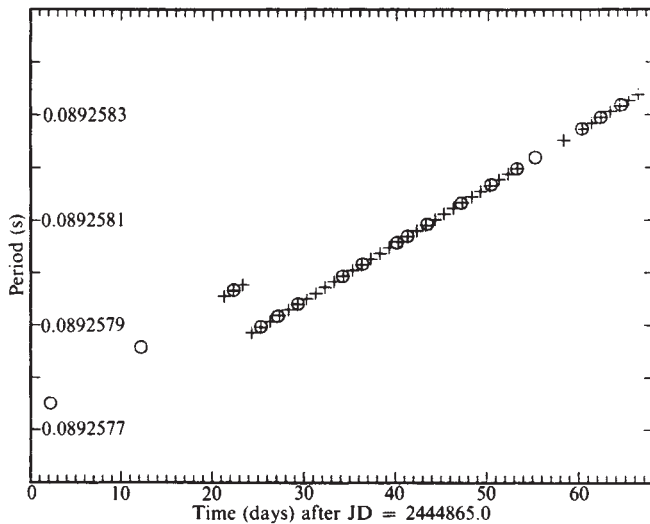


Fig. 1 Mean period of the Vela pulsar from each day's observations at Hobart and Tidbinbilla.

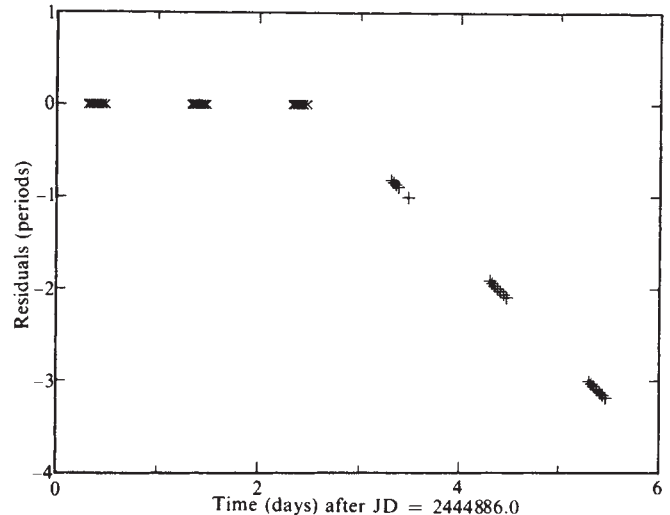


Fig. 2 Phase residuals from fitting the pre-jump timing parameters to the Hobart data for the three days preceding and following the jump.

techniques. In this calculation the optical position⁴ with the *E*-terms of annual aberration removed, RA (1950) 08 h 33 min 39.27 s, Dec. (1950) $-45^{\circ}00'10.2''$, is assumed.

Periods derived from observations at Hobart and Tidbinbilla between 19 September 1981 and 22 November 1981 are plotted in Fig. 1, and show the usual secular increase of about 10.7 ns per day together with a decrease of 102.2 ns between observing sessions on 10 October and 11 October. This is the fifth major period jump observed in the period of the Vela pulsar. Its characteristics are similar to those of previously observed jumps except that the magnitude of the jump is smaller. The magnitudes of previously observed jumps are⁵ 208 ns (1969), 174 ns (1971), 179 ns (1975) and 272 ns (1978).

Pre-jump arrival time data from Tidbinbilla for July to October 1981 are well fitted by a cubic giving the pulse phase at time t as

$$\phi(t) = \phi_0 + \nu(t - T_0) + \frac{1}{2}\ddot{\nu}(t - T_0)^2 + \frac{1}{6}\ddot{\nu}(t - T_0)^3$$

The parameters of this fit are given in Table 1.

Figure 2 shows the residuals from this fit for Hobart data from 8–13 October. The residuals for 11–13 October show a linear departure from the predicted phase, indicating the presence of a period jump. Simple extrapolations suggest that the jump occurred shortly after observations finished on 10 October, that is at about 0315 UT on 10 October 1981 (JD = 2444888.57). This is the first Vela jump in which the interval between the pre- and -post-jump observations is sufficiently small to permit a unique determination of the jump epoch.

The post-jump recovery has commonly been modelled in terms of the two-component neutron-star model² in which a certain fraction of the initial frequency jump decays exponentially. In this model the pulsar frequency takes the form for $t \geq T_j$

$$\nu(t) = \nu_p(t) + \Delta\nu_c + \Delta\nu e^{-(t-T_j)/\tau}$$

where T_j is the jump epoch, $\nu_p(t)$ is the extrapolated pre-jump frequency, $\Delta\nu_c$ is the permanent part of the frequency jump and $\Delta\nu$ is the part which subsequently decays with time constant τ . The parameter Q representing the fraction of the initial jump which decays is then given by

$$Q = \Delta\nu / (\Delta\nu_c + \Delta\nu)$$

This function is not a good fit to the data unless the points for the 9 days immediately after the jump are excluded. The residuals from such a fit to the Hobart data are shown in Fig. 3a. The departure of the observed residuals from this function is well described by an additional exponential term with a short time constant. Figure 3b shows the residuals when the Hobart data are fitted with the function

$$\nu(t) = \nu_p(t) + \Delta\nu_c + \Delta\nu_1 e^{-(t-T_j)/\tau_1} + \Delta\nu_2 e^{-(t-T_j)/\tau_2}$$

Table 1 gives the best fit values for the parameters. The quoted uncertainties are twice the formal standard errors of the least-squares fit. The random nature of the residuals in Fig. 3b shows that, down to the level of these residuals, the sum of two exponentials describes the period jump completely. Of the total frequency jump

$$\begin{aligned} \Delta\nu_j &= \Delta\nu_c + \Delta\nu_1 + \Delta\nu_2 \\ &= (1.2828 \pm 0.0010) \times 10^{-5} \text{ Hz} \end{aligned}$$

0.7% and 17.6% decayed away with the short and long time constants respectively. Table 1 summarizes the post-jump decay parameters.

We have modelled the long-term recovery from the jump with the exponential function required by previous jumps. However, note that the present case, with only 40 days of data, could be described equally well by perturbations in the coefficients of the cubic used to describe the pre-jump behaviour.

No other parameters of the pulsar changed at the time of the period jump. The pulse profiles before and after the jump were compared for a variety of integration times and in all cases they agree down to the noise level. Some polarization information is available as we record the two linear polarizations separately; no significant change was seen at the time of the jump.

It is unfortunate that the initial jump in frequency was not observed. However, the first post-jump observations were <18 h after the estimated jump epoch, much less than for any previous Vela jump. It is significant that the first post-jump observations show no evidence for continued spin-up of the pulsar. This limits the duration of the initial spin-up to <18 h. Greenstein⁶, for example, has suggested that in some circum-

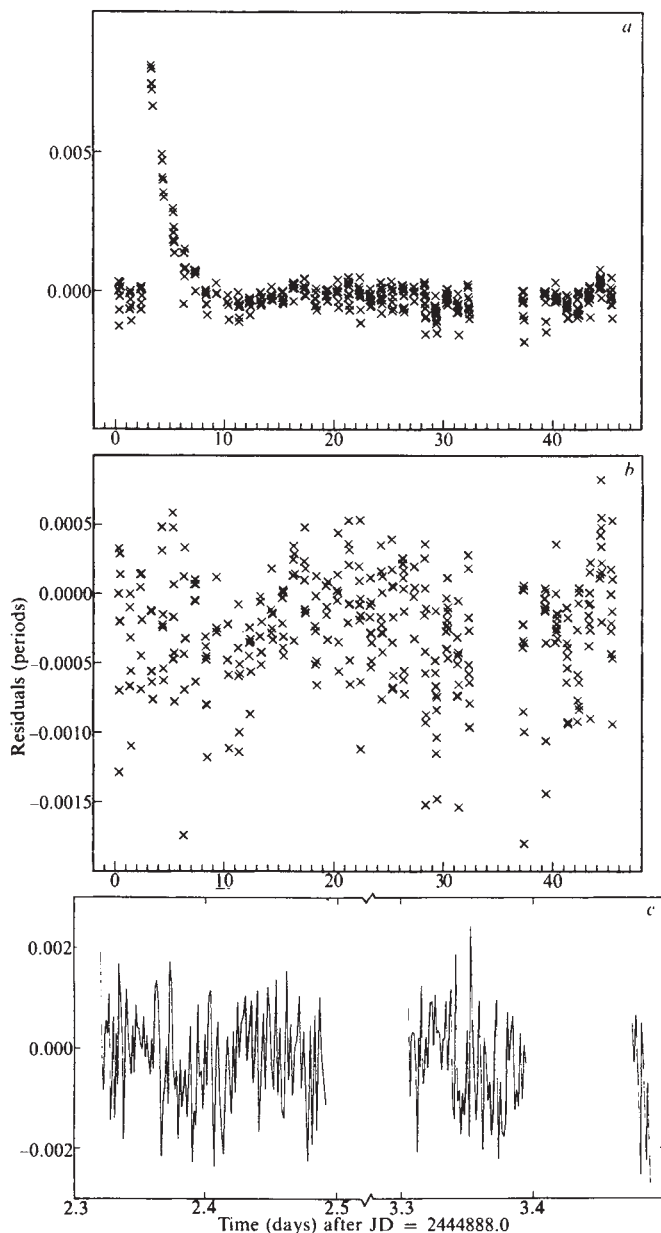


Fig. 3 Residuals from fitting the expressions given in the text to the Hobart data. *a*, Results from fitting a single exponential to data excluding those points for days 2444889–897; each point represents 30 min of observing. *b*, Results from fitting the sum of two exponentials; each data point represents 30 min of data. *c*, An expanded version of *b* for the days either side of the glitch. Each point represents a 1,000-period integration; successive points are joined by straight lines to show the time sequence.

stances the initial spin-up could take a significant time to occur. The initial short-term decay reported here is superficially similar to the rapid decays following previous events reported by Downs⁵, but the time scales are quite different. Decay times reported for the earlier events are in the range 40–80 days, comparable with the total data span of the present observations. On this time scale, and excluding the immediate post-jump data, the present observations are fitted extremely well by the standard two-component model with parameters similar to those derived for previous events, such as $Q \sim 0.2$ and τ a few hundred days (see Fig. 3c). It remains to be seen if this exponential decay continues on longer time scales or if the value of ν stabilizes at some value as observed by Downs in previous events.

The observed rapid initial decay may indicate the presence of a third component in the neutron star. By analogy with the two-component model², the relatively small value of $\Delta\nu_1/\nu$ (8.2×10^{-9}) and the very short decay time (1.6 days) suggest that this third component and the neutron star crust (to which the pulses are assumed to be locked) are closely coupled. The relative moment of inertia of the crust and the third component (normally obtained from the Q value) cannot be estimated, as it is not known what fraction of the derived $\Delta\nu_c$ is associated with the short-term decay. Note that this short-term feature resembles the glitches (sudden, discontinuous changes in frequency) observed in the Crab Nebula pulsar^{7,8} in the magnitudes of $\Delta\nu/\nu$ and τ .

Greenstein⁶ has proposed a model for the Vela period jumps based on the response of a two-component neutron star to a temperature perturbation. For the case computed (a 50% jump in the temperature), the fractional change in pulse frequency, $\Delta\nu/\nu$, is of the order observed ($\sim 10^{-6}$), but the predicted change in frequency derivative, $\Delta\dot{\nu}/\dot{\nu}$, after ~ 1 day is about 2.5, far greater than the observed value of 0.04. It is possible that with variation of the assumed parameters the computed value of $\Delta\dot{\nu}/\dot{\nu}$ in Greenstein's model could be reduced. An alternative model in which the spin-up results from the sudden unpinning from the neutron star crust of vortices in the neutron superfluid has been proposed by Alpar *et al.*⁹. This model appears capable of accommodating the various spin-up and decay time scales observed in the past, but does not specifically predict the rapid initial decay of the present observations.

Greenstein¹⁰ has also predicted the behaviour of $\Delta\dot{\nu}/\dot{\nu}$ after an abrupt increase in $\Delta\nu/\nu$ of 10^{-6} . His Fig. 3 shows $\Delta\dot{\nu}/\dot{\nu}$ as two linear segments with a sharp change in slope between them on a semi-log plot. Such behaviour should be closely modelled by the sum of two exponentials of the type we have fitted, although the parameters of our fit are quite different from those predicted by Greenstein.

Downs⁵ suggested on the basis of the four previously observed jumps that a correlation existed between the value of $\dot{P}$ (or $\ddot{\nu}$) before a given jump and the magnitude ΔP (or $\Delta\nu$) of the jump. Downs quotes a value of $\dot{P} = (-5.5 \pm 1.0) \times 10^{-24} \text{ s s}^{-2}$ for the interval after the 1978 jump to the end of 1980. This would lead to a predicted $\Delta P \sim 180$ ns, considerably larger than the observed value of 102 ns. However, for the pre-jump fit quoted in Table 1, $\dot{P} = -\ddot{\nu}/\nu^2 = (-3.1 \pm 0.4) \times 10^{-24} \text{ s s}^{-2}$ leading to a predicted $\Delta P \sim 90$ ns, close to the observed value, given the uncertainties. The fact that the derived $\dot{P}$ for the latter part of 1981 is less than the value quoted by Downs for 1979–80 suggests that for this event the value of $\dot{P}$ did not stabilize but continued to approach zero up to the time of the next jump.

It has been suggested that immediately preceding and following a glitch the pulsar period may vary erratically. We have examined the data for the transits either side of the glitch and see no evidence of any such effects. Figure 3c shows a plot of the residuals from fitting the sum of two exponentials to the original 1,000-pulse integrations obtained on these days. The data points have been joined by straight line segments so as to make their order more apparent. The r.m.s. of the fit on these days is essentially the same as for all other days.

Received 26 October 1982; accepted 5 January 1983.

1. McCulloch, P. M., Hamilton, P. A., Royle, G. W. R. & Manchester, R. N. *IAU Circ. No.* 3644 (1981).
2. Baym, G., Pethick, C., Pines, D. & Ruderman, M. *Nature* **224**, 872–874 (1969).
3. Manchester, R. N., Hamilton, P. A., Goss, W. M. & Newton, L. M. *Proc. Astr. Soc. Aust.* **3**, 81–83 (1976).
4. Manchester, R. N. *et al. Mon. Not. R. astr. Soc.* **184**, 159–170 (1978).
5. Downs, G. S. *Astrophys. J.* **249**, 687–697 (1981).
6. Greenstein, G. *Astrophys. J.* **231**, 880–895 (1979).
7. Boynton, P. E. *et al. Astrophys. J.* **175**, 217–241 (1972).
8. Lohson, E. *Nature* **258**, 688–689 (1975).
9. Alpar, M. A., Anderson, P. W., Pines, D. & Shaham, J. *Astrophys. J. Lett.* **249**, L29–L33 (1981).
10. Greenstein, G. *Astrophys. J.* **208**, 836–848 (1976).

D/H ratios of coals and the palaeolatitude of their deposition

J. W. Smith*, D. Rigby*, P. W. Schmidt†
& D. A. Clark†

* CSIRO Division of Fossil Fuels and † CSIRO Division of Mineral Physics, PO Box 126, North Ryde, New South Wales 2113, Australia

Various interpretations of D/H ratios in coals have been suggested¹⁻⁴. The factors most likely to influence these ratios seem to be: (1) prevailing climatic conditions and the related deuterium contents of available fresh water (precipitation) during plant growth; and (2) the fractionation of stable hydrogen isotopes during both plant growth and subsequent sediment maturation. On theoretical grounds, the global distribution of the hydrogen isotopes in precipitation is expected to favour an increased deuterium concentration at the Equator relative to the poles. Published measurements of the D/H ratio in present day precipitation at different latitudes, as shown in Fig. 1⁵, follow this trend. Clearly the well-established global variation in the isotopic composition of precipitation with regard to both climate and latitude provides a reasonable explanation for the differences observed in D/H ratios for coal from different continents as has been proposed for coals from Germany and South Africa¹, and for isotopic variations seen in peat². The D/H ratios for Australian coals reported here are consistent with their respective latitudes of deposition predicted from continental drift and hence the implied dependence on palaeoclimate.

Other authors³ conclude that variations in the D/H ratios in coals are primary source-controlled (regardless of their stage of thermal evolution) and infer that differences in contributing plant associations are responsible for the variations observed rather than maturational effects. In this latter respect, during the catagenesis of coals beyond the sub-bituminous stage, the main vehicle for hydrogen loss is methane. Measurements on over 100 related seam gas and bituminous coal samples from Australian mines have shown methane generally to be depleted in deuterium relative to the parent coal by 70–100%. Therefore individual coals are likely to undergo well-defined changes in isotopic composition during maturation. However, if maturation were the only mechanism to determine the isotopic composition of coals in general, δD values of coals would be expected to increase with increasing maturity. In fact the reverse is seen for the coals examined here. We regard this inverse correlation as further evidence that the wide range of isotopic

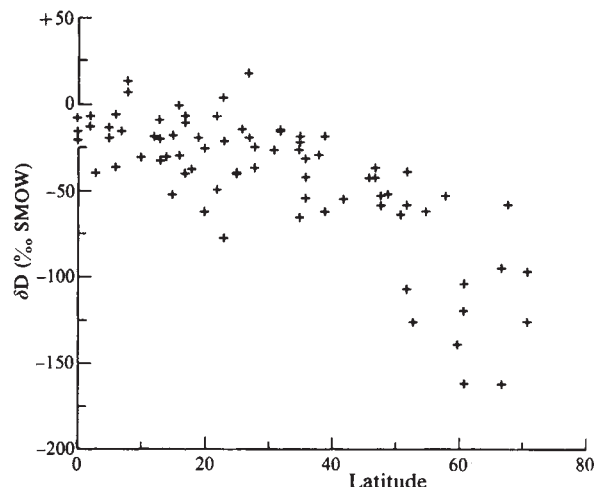


Fig. 1 D/H ratios in precipitation (after ref. 5).

variations exhibited by coals is not primarily a result of maturation reactions.

In contrast, the relatively higher concentration of exchangeable protons in brown coals from Bass Strait than in bituminous coals from the Bowen Basin has been suggested⁴ as a cause of the major isotopic differences seen between these Australian samples. As yet none of these suggestions alone seems to account satisfactorily for the observed distribution of hydrogen isotopes in Australian coals.

Before any meaningful comparison can be drawn between the D/H ratios in coals and the freshwater (precipitation) available for original plant growth, an estimation of the fractionation of hydrogen isotopes accompanying plant growth is required. Fractionations due to this process for whole plants range from –30% to –55% (refs 1, 6) (δD values are quoted in ‰ relative to standard mean ocean water (SMOW)). However, quite different fractionations are reported for specific plant components. For example, at 10 °C the fractionation between tree cellulose and available water is only 12% (ref. 7), whereas a larger fractionation of 92% accompanies lipid synthesis from carbohydrates⁸. Neither of these extreme isotopic fractionations significantly affects the weighted isotopic composition of whole coals. Cellulose is largely and preferentially decomposed in the earliest biological stages of coal formation and the coal lipids, although exhibiting a major depletion in deuterium content⁸, are present in minor concentrations. Therefore an average isotopic fractionation of the stable isotopes of hydrogen between precipitation and coals of approximately –40% might be expected if further isotopic fractionation during continued maturation of plant residues is minimal.

Table 1 Location, age and palaeolatitude of coal samples from Australia and Antarctica

Sample no.	Location	Carbon (% dry ash-free)	Hydrogen (% dry ash-free)	Lat. (°S)	Long. (°E)	δD^* (‰ SMOW)	Age (Myr)	Palaeolatitude (deg.)
1	Collinsville	87.2	4.7	21	148	–134	230–260	62–81
2	Styx	83.4	5.0	23	150	–132	90–110	53–66
3	Portland	85.5	5.5	25	153	–140	90–110	55–68
4	Lagoon	82.5	5.5	28	153	–142	180–200	56–66
5	Bluff	81.5	5.4	28	153	–149	180–200	56–66
6	Millmerran	78.6	6.6	28	151	–159	150–170	75–60
7	Nymboida	86.4	5.5	30	153	–168	180–200	58–68
8	Leigh Creek	73.3	4.1	30	138	–112	170–200	48–58
9	Lake Phillipson	73.4	4.6	31	154	–133	250–270	67–79
10	Yallourn	72.1	7.3	38	146	–100	10–15	41–43
11	Traralgon	65.9	4.7	38	146	–94	35–40	50–54
12	Loy Yang	68.6	4.7	38	147	–110	15–30	43–49
13	Wonthaggi	80.2	5.3	39	146	–130	115–130	67–75
14	Antarctica (Beaver Lake)	77.1	5.3	71	68	–139	230–250	68–53

* Standard deviation = $\pm 1\%$.

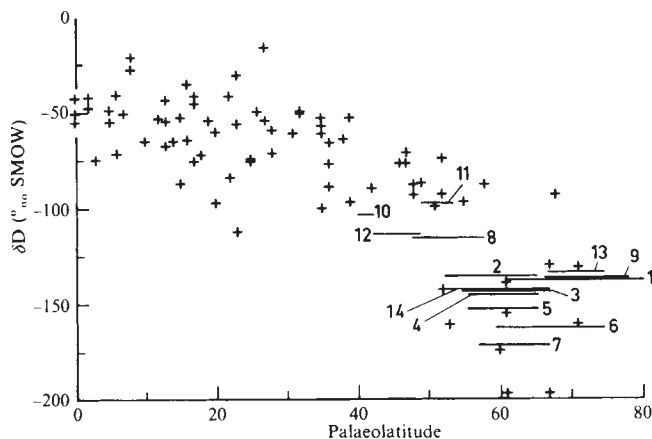


Fig. 2 Correlation between measured D/H ratios in coals (—) and calculated D/H ratios in land plants (+) at corresponding palaeolatitudes.

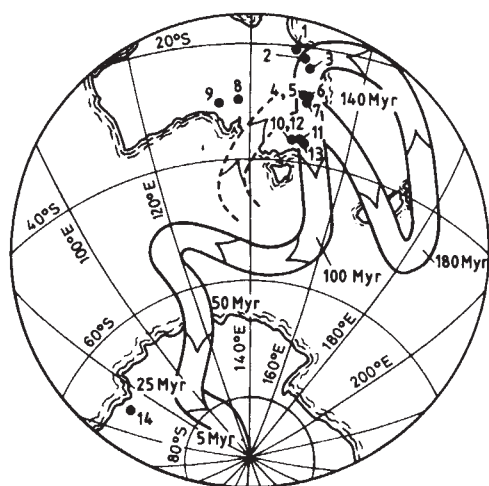


Fig. 3 Sample locations and South Pole track relative to Australia (after ref. 9).

Accordingly, if a reduction of 40% is applied to the data on precipitation presented in Fig. 1, a general relationship between D/H ratio in land plants and their latitudes of growth results (Fig. 2).

A comparison of the present latitudes and δD values of some southern coals (Table 1) reveals no simple relationship. However, the ages of the coals are such that a valid comparison of latitude and δD values must allow for the effects of continental drift. A recent compilation⁹ of the apparent polar wander path (APWP) data for Australia since the late Palaeozoic is shown in Fig. 3. Clearly the present latitudes of the coal deposits are very different from their inferred palaeolatitudes.

The palaeolatitudes of the coals listed in Table 1 correspond to the great arc lengths in degrees between the present geographical coordinates and the positions on the APWP corresponding to the ages of the coals. The Antarctic locality was first transformed to its pre-drift position with respect to Australia by a rotation of 30.9° about an Euler pole at 11.9°N, 30.8°E¹⁰. Because a time interval is given for the formation of each coal deposit, a corresponding range of palaeolatitudes has been calculated. The maximum errors in estimating the latitudes of the coals during deposition are represented in Fig. 2 by the error bars. When the data are presented in this form it is apparent that a correlation exists between D/H ratios in these coals and their palaeolatitudes of formation. Further, these D/H ratios, as a function of palaeolatitude, compare favourably with calculated D/H ratios in land plants and their respective latitudes of growth.

These findings, which are based on observations on coals deposited in higher palaeolatitudes, suggest that little hydrogen isotope change accompanies the coalification of bulk land plant material. Therefore, D/H ratios in coals reflect predominantly the isotopic composition of the available freshwater during plant growth.

Received 24 August 1982; accepted 11 January 1983.

1. Schiegl, W. E. & Vogel, J. C. *Earth planet. Sci. Lett.* **7**, 307–313 (1970).
2. Schiegl, W. E. *Science* **175**, 512–513 (1972).
3. Redding, C. E., Schoell, M., Monin, J. C. & Durand, B. in *Advances in Organic Geochemistry, 1979* (eds Douglas, A. G. & Maxwell, J. R.) 711–723 (Pergamon, Oxford, 1980).
4. Smith, J. W., Gould, K. W. & Rigby, D. *Org. Geochem.* **3**, 111–131 (1982).
5. Dansgaard, W. *Tellus* **16**, 436–468 (1964).
6. Smith, B. N. & Epstein, S. *Pl. Physiol.* **46**, 738–742 (1970).
7. Yapp, J. C. & Epstein, S. *Nature* **297**, 636–639 (1982).
8. Rigby, D., Batts, B. D. & Smith, J. W. *Org. Geochem.* **3**, 29–36 (1981).
9. Emberton, B. J. J. in *Paleoreconstruction of the Continents. Geodynamics Series* **2**, 77–92 (American Geophysical Union, 1981).
10. Norton, I. D. & Sclater, J. G. *J. geophys. Res.* **84**, 6803–6830 (1979).

Sulphur isotope abundances in Aphebian clastic rocks: implications for the coeval atmosphere

Keiko Hattori*, Finley A. Campbell* & H. Roy Krouse†

* Department of Geology and Geophysics, and † Department of Physics, The University of Calgary, Calgary, Alberta, Canada T2N 1N4

The evolution of the atmosphere has been of interest to many scientists because it relates to the history of the ocean, the sedimentary cycle, the geochemical cycle of volatile elements, and the development of life. It also has a key role in the genesis of iron formations and gold and uranium deposits. Although it is generally accepted that the primitive atmosphere was reducing, a debatable aspect is when and how the oxygen content increased. A rise of atmospheric oxygen pressure about 2,000 Myr BP is suggested by the appearance of redbeds and the banded iron formations^{1,2}. On the other hand, high atmospheric oxygen in very early Precambrian time, as a result of photodissociation of water, has been proposed³. This is supported by occurrences in Precambrian rocks of sulphates^{4,5}, volatile elements⁶, detrital minerals^{5,7} and palaeosols⁸, which are similar to those found in younger rocks. Because the change of sulphur valency is often accompanied by isotopic fractionation, sulphur isotope abundances in minerals have been used to infer the environment at the time of their formation. It occurred to us that additional sulphur isotope analyses of pyrite from continental sedimentary rocks might provide evidence relative to the question of oxygen development in the atmosphere. Accordingly, a major study of sulphur isotope abundances in pyrite from fluvial sedimentary rocks in the Elliot Lake area, Canada, and for comparison, data from occurrences in Northwest Territories, Canada, Witwatersrand, South Africa, and Jacobina, Brazil, are now reported and implications respecting the coeval atmosphere discussed.

The Huronian Supergroup of early Aphebian age occurring in the Elliot Lake area, Ontario, was deposited between ~2,200 and ~2,500 Myr BP based on ages of intrusives in the sedimentary rocks^{9,10} and the latest granite in Archaean terrain (see Table 1 and refs 9, 10). The Matinenda formation, the base of the Supergroup, is ~2,350-Myr old as determined from the age of the Copper Cliff Rhyolite which is intercalated with the formation south of Sudbury¹¹. The metamorphic grade is greenschist, and the original sedimentary features remain¹². There is no evidence of intense hydrothermal activity nor substantial recrystallization of minerals in the study area¹². This essentially

Table 1 Stratigraphical succession of the Huronian Supergroup

Palaeozoic formations	
Nipissing Diabase	2,160 Myr BP
Cobalt Group	
Quirke Lake Group	
Hough Lake Group	
Elliott Lake Group	McKim formation
	Matinenda formation
Archaean basement rocks	
	>2,550 Myr BP

unaltered character and the continental origin provide a good sequence in which to evaluate sulphur isotope patterns and atmospheric conditions at the time of formation.

The Matinenda formation consists of mineralogically mature quartzite, which occupies over 90% of the total volume, and 10% of quartz-pebble conglomerate. The areal distribution and textural features suggest that the sediments were deposited in braided streams^{12,13}. The conglomerate beds represent the channels in the stream and contain heavy radioactive minerals which constitute the largest uranium resource in North America.

Quartzites consist of well-rounded and well-sorted quartz grains, sericite, and small amounts of feldspar and pyrite (~0.001 to ~0.1 wt %). The conglomerates consist of quartz-pebbles, sericite, pyrite (10–20 wt %), feldspar, and minor amounts of heavy minerals.

Grain size analyses of the conglomerate show that there is a good correlation between the maximum size of pyrite, uraninite, and quartz-pebbles in a sample, indicating that the coarse euhedral to semihedral pyrite grains were deposited by hydraulic processes¹³. Earlier petrographic studies^{12,14} as well as recent microprobe analyses¹⁵ support the view that some uraninite and pyrite are detrital. Textures of some fine anhedral pyrite and overgrowths on euhedral pyrite in the conglomerate also suggest that detrital as well as chemically formed pyrite is present^{12,16}. Pyrites in the quartzite are anhedral and occur interstitially to the clasts, indicating that they are not detrital. The lack of hydraulic equivalence of the size of pyrite in the quartzite suggests formation during diagenesis. Because pyrite tends to retain its original sulphur isotope composition during the process of recrystallization, the measured $\delta^{34}\text{S}$ values represent those of either detrital pyrite or pyrite which was formed from solution.

Pyrite in powdered rock samples was oxidized using $\text{HNO}_3\text{--Br}_2$. The sulphate formed was then precipitated as BaSO_4 by adding BaCl_2 solution. Extraction of SO_2 gas for mass spectrometric analysis was performed by direct thermal decomposition using NaPO_3 (ref. 17) or by oxidation of Ag_2S (ref. 18). The isotopic analyses were performed with a mass spectrometer based on Micromass 602 components. The data are expressed with respect to the Cañon Diablo meteorite standard using the familiar δ notation and standard deviations from the means are given (1σ).

The $\delta^{34}\text{S}$ values of pyrite from conglomerate and quartzite in the Matinenda formation from Quirke II Mine are presented in Fig. 1, together with a schematic section. The 44 samples represent complete coverage of the section (~65 m). Their $\delta^{34}\text{S}$ values ranging from -1.0 to $+1.1\%$ ($+0.21 \pm 0.54\%$) are close to the meteorite standard, independent of rock types: $+0.20 \pm 0.55\%$ for quartzite and $+0.14 \pm 0.57\%$ for conglomerate. Values near 0‰ are also found in quartzite and conglomerate of the same formation in the Rio Algom Mine, about 15 km south of the Quirke II Mine (Table 2). Euhedral coarse pyrite crystals of detrital origin in the conglomerate were manually separated and their $\delta^{34}\text{S}$ values were also found to be close to 0‰ (Table 3).

Pyrite of diagenetic and authigenic origin in the conglomerate has a finer grain size than the detrital pyrite. $\delta^{34}\text{S}$ values for different size fractions were determined (Table 3). Since pyrite which was formed from solution during early diagenesis occurs

Table 2 Sulphur isotope values of pyrite

		$\delta^{34}\text{S}_{\text{CD}}$ (‰)
Rio Algom Mine, Elliot Lake area		
E 650	Quartzite, 1–3 ft above the main uraniferous conglomerate	+1.1
E 651	Typical uraniferous quartz-pebble conglomerate	–0.1
E 649	Quartzite, base of the Matinenda formation	+0.4
Denison Mine, Elliot Lake area		
81930-1-2-P	Disseminated in graphitic matter	–0.6
81930-1-2-P	Adjacent to graphitic matter	–0.5
81930-1-2-P	Adjacent to graphitic matter	+0.2
Panel Mine, Elliot Lake area		
RAD 14-11	Sandstone, below the organic seam	+0.1
RAD 14-11	Adjacent to stratified sulphur-rich organic matter	–0.2
RAD 14-11	Disseminated in the organic seam	+1.1
Carbon Leader Reef, South Africa		
82813-5	3 Mudballs	+3.9
82813-6	7 Mudballs	+4.0
82813-7	5 Mudballs	0.0
82825-1	7 Mudballs	+2.6
82826-1	5 Mudballs	–0.5
82813-13	Disseminated in graphitic matter	+4.0
81813-14	Disseminated in graphitic matter	+3.5
President Mine, South Africa		
81-9-9-1	Uraniferous/quartz-pebble conglomerate	+1.8
81-9-9-2	Uraniferous/quartz-pebble conglomerate	+2.4
81-9-9-3A	Uraniferous/quartz-pebble conglomerate	+1.7
81-9-9-3C	Uraniferous/quartz-pebble conglomerate	+1.9
Padlei Lake area, North-west Territories		
GY 69-137-A-A	Quartz-pebble conglomerate	–1.3
GY 69-137-A-B	Quartz-pebble conglomerate	–1.7
GY 69-137-D-A	Quartz-pebble conglomerate	–2.0
GY 69-137-D-B	Quartz-pebble conglomerate	–1.5
Jacobina, Brazil		
16680	Base of Piritose Reef, uraniferous quartz-pebble conglomerate	+0.8
16683	Liberina Reef, quartz-pebble conglomerate	–0.1

on the outer rim of grains, each size fraction of pyrite was separated into outer rim and inner core portions by an etching technique using $\text{HNO}_3\text{--Br}_2$. There was no noticeable difference between values for the rim of $+0.14 \pm 0.21\%$ and the core $+0.16 \pm 0.27\%$.

Organic matter in sediments has an important role in the fixation of sulphide by providing a local reducing environment. Reduction of sulphate can produce variations in the isotopic compositions of the resultant sulphides. $\delta^{34}\text{S}$ values were determined for pyrite closely associated with as well as occurring in organic matter in the Matinenda formation in the Denison Mine and Pronto Mine (Table 2) to look for variations. No significant differences in the sulphur isotopic compositions of pyrite from organic-rich and carbon-free environments were found.

$\delta^{34}\text{S}$ values of 75 samples from the Matinenda formation consisting of detrital pyrite, pyrite formed from solution and formed in a reduced environment are all near 0‰. These contrast with pyrite in modern sediments (-42 to $+30\%$ (refs 19–21)), uraniferous continental sandstone of Mesozoic and Tertiary age (-50 to $+29\%$ (refs 22, 23) and Cretaceous braided stream sedimentary rocks ($+1$ to $+34\%$; unpublished data).

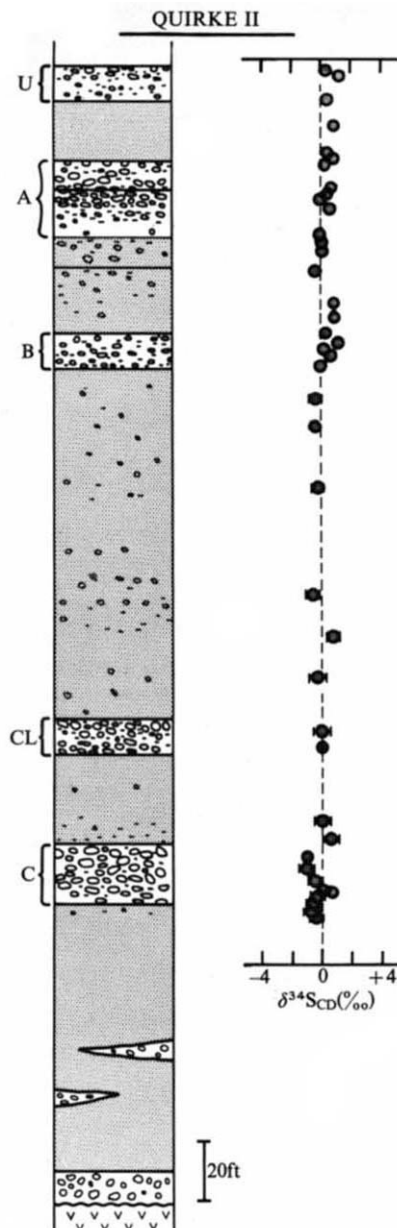


Fig. 1 Sulphur isotope ratios in pyrite from quartz-pebble conglomerate and the quartzite in the New Quirke II mine. Main conglomerate beds are labelled U, A, B, CL and C.

The values of $\sim 0\%$ cannot be attributed to the effect of metamorphism nor secondary isotope exchange processes because large $\delta^{34}\text{S}$ variations in pyrite have been preserved in highly metamorphosed sedimentary rocks. For example, the Michipicoten and Woman River Iron formations of Algoma type in Ontario which were affected by strong tectonic deformation and regional metamorphism during the Kenoran orogeny display $>14\%$ variations in $\delta^{34}\text{S}$ values within 6 m (refs 24, 25). Another example is the Homestake gold mine (ref. 26). Therefore, the observed $\delta^{34}\text{S}$ values in the Matinenda formation represent the original values when pyrite was formed.

Because the wide variations in $\delta^{34}\text{S}$ values in Phanerozoic sedimentary rocks are attributed to kinetic isotope effects during the reduction of sulphate, the values $\sim 0\%$ in the whole section of Matinenda formation suggest that sulphate reduction did not have a significant role in the formation of pyrite. Theoretically, complete reduction of sulphate accompanied by thorough mixing of the product sulphide could have produced similar results. However, such conditions throughout an entire geological section of alternating conglomerates and quartzites in an open fluvial sedimentary system are highly unlikely.

Table 3 $\delta^{34}\text{S}$ values of pyrite of different grain size in the uraniferous conglomerate ($\%$)

Grain size*	Measured part†	E 651‡	E 646§	E 659	E 660¶
E		-1.5	-0.1	+0.1	ND
>no. 60	R	ND	+0.3	+0.1	ND
	C	+0.3	+0.2	0.0	-0.2
no. 60 < >no. 80	R	+0.3	+0.1	-0.2	ND
	C	+0.2	0.0	+0.4	+0.1
no. 80 < >no. 150	R	+0.2	+0.4	-0.1	ND
	C	+0.3	+0.7	-0.2	ND
<no. 150		-0.6	-0.7	+0.2	+0.1

* E, semihedral grains ($\phi > 3$ mm).

† R, rim fraction; C, core fraction. The two fractions are chemically separated (see text).

‡ Sample from the Rio Algom Mine.

§ Sample from the Panel Mine.

|| Sample from the Denison Mine.

¶ Sample from the Stanrock Mine.

The interpretation that the pyrite was not the product of sulphate reduction is further reinforced by $\delta^{34}\text{S}$ values $\sim 0\%$ for pyrite intimately associated with the organic matter (Table 2). Because $\delta^{34}\text{S}$ values for pyrite associated with the organic matter did not differ from those of pyrite in carbon-free regions, pyrite formation by reduction of sulphate is unlikely. This, in turn, suggests that sulphate was absent at the time of pyrite formation.

The Matinenda formation is the product of intense weathering in warm humid conditions²⁷. The detritus was transported in an aerated stream environment, and deposition involved repeated reworking and periods of exposure to the atmosphere. If the atmosphere contained appreciable oxygen, pyrite and dissolved sulphur species would have been oxidized to sulphate ion. Subsequent reduction of the sulphate would have produced pyrite with widely varying $\delta^{34}\text{S}$ values.

The above interpretation contradicts the suggestions that pyrite in the Matinenda formation is due to the activity of sulphate reducing bacteria^{28,29} or sulphate reduction in groundwater³⁰. Furthermore, the $\delta^{34}\text{S}$ variations in the Phanerozoic uraniferous sandstones, due to oxidation-reduction reactions²², indicate the different processes of pyrite formation in the two types of deposits.

The $\delta^{34}\text{S}$ values $\sim 0\%$ for core portions of detrital and euhedral pyrite suggest that the source is probably Archaean basement rocks. Dissolved HS^- and H_2S in streams or groundwater may have combined with Fe^{2+} to form pyrite. Thus, $\delta^{34}\text{S}$ values $\sim 0\%$ in Archaean source rocks were preserved during the sedimentary cycle and appear in the Proterozoic fluvial sedimentary rocks.

Our conclusion of low atmospheric oxygen $\sim 2,300$ Myr BP implies that other continental sedimentary rocks of similar or older age should have sulphur isotope values close to 0% . Supporting evidence is found in various localities. Thode *et al.*³¹ reported $\delta^{34}\text{S}$ values of -0.5 to $+0.1\%$ for sedimentary rocks of similar age and possibly fluvial origin south of the Sudbury basin. $\delta^{34}\text{S}$ values of $+0.1$ to $+4.0\%$ were determined for pyrite^{32,33} in $\sim 2,700$ -Myr old quartz-pebble conglomerate of braided stream origin in the Witwatersrand basin, South Africa^{34,35}. We found similar values for pyrite from the President Mine, Orange Free State and the Carbon Leader Reef (Table 2). In the latter deposit, 'mudball pyrite', which is thought to have been formed on alluvial fans³⁶ has $\delta^{34}\text{S}$ values of -0.5 to $+4.0\%$. Pyrite in graphitic fossil plant from the same locality^{37,38} has $\delta^{34}\text{S}$ values $+3.5$ and $+4.0\%$ (Table 2). Pyrite in quartz-pebble conglomerates in the Montgomery formation from the Padlei Lake area, North-west Territories, Canada, of early Proterozoic age³⁹ give values to 0% (-2.0 to -1.3% (Table 2)). Thus the $\delta^{34}\text{S}$ values for pyrite from various localities in early Proterozoic continental sedimentary rocks are best attributed to global anaerobic conditions which inhibited

oxidation of sulphur in the sedimentary cycle. This conclusion supports the models of low atmospheric oxygen during this time as proposed by Cloud¹ and Garrels *et al.*²

This study was supported by NSERC grants A-8176 to H.R.K. and A-1042 to F.A.C. We thank the staff of the Quirke Mine and the Denison Mine for permitting sample collection. A. Kidwell, V. Ruzicka, G. M. Young, N. D. Smith and G. D. Williams donated samples from the President Mine, the Panel Mine, Padlei Lake area and Jacobina, Brazil, the Carbon Leader Reef and Mannville Quartzite, Alberta, respectively.

Received 15 November 1982; accepted 25 January 1983.

- Cloud, P. E. *Jr Science* **160**, 729–735 (1968).
- Garrels, R. M., Perry, E. A., Jr & McKenzie, F. T. *Econ. Geol.* **68**, 1173–1179 (1973).
- Brinkman, R. T. *J. geophys. Res.* **74**, 5355–5368 (1969).
- Vinogradov, V. I., Reimer, T. O., Leites, A. M. & Smelov, S. B. *Lithol. Pol. Isk.* **4**, 12–27 (1976).
- Clemmey, H. & Badham, N. *Geology* **10**, 141–146 (1982).
- Dimroth, E. & Kimberley, M. M. *Can. J. Earth Sci.* **13**, 1161–1185 (1976).
- Simpson, P. R. & Bowles, J. F. W. *Phil. Trans. R. Soc. A* **286**, 527–548 (1977).
- Gay, A. L. & Grandstaff, D. E. *Precamb. Res.* **12**, 349–373 (1980).
- Van Schmus, W. R. *J. Geol.* **73**, 755–780 (1965).
- Fairbairn, H. W., Hurley, P. M., Card, K. D. & Knight, C. V. *Can. J. Earth Sci.* **6**, 489–497 (1969).
- Roscoe, S. M. *Geol. Ass. Can. Spec. Pap.* **12-271**, 31–47 (1973).
- Roscoe, S. M. *Geol. Surv. Can. Pap.* **68-40** (1969).
- Theis, N. J. *Bull. geol. Surv. Can.* 304 (1979).
- Ramdohr, P. *Geol. Soc. S. Afr.* **71**, 67–100 (1958).
- Meddaugh, W. S. & Holland, H. D. *Geol. Soc. Am. Abst. Prog.* **13**, 481 (1980); **14**, 509 (1981).
- Robinson, A. G. & Spooner, E. T. *Nature* **299**, 622–624 (1982).
- Halas, S. & Wołacewicz, W. P. *Analyt. Chem.* **63**, 686–689 (1981).
- Robinson, B. W. & Kusakaba, M. *Analyt. Chem.* **47**, 1179–1181 (1975).
- Kaplan, I. R., Emery, K. O. & Rittenberg, S. C. *Geochim. cosmochim. Acta* **27**, 297–331 (1963).
- Schwarcz, H. P. & Burnie, S. W. *Miner. Depos.* **19**, 264–277 (1973).
- Chambers, L. A. *Geochim. cosmochim. Acta* **46**, 721–728 (1982).
- Jensen, L. *Econ. Geol.* **53**, 598–616 (1958).
- Goldhaber, M. B., Reynolds, R. L. & Rye, R. O. *Econ. Geol.* **73**, 1690–1705 (1978).
- Goodwin, A. M., Monster, J. & Thode, H. G. *Econ. Geol.* **71**, 870–891 (1976).
- Thode, H. G. & Goodwin, A. M. *Precamb. Res.* (in the press).
- Rye, D. M. & Rye, R. O. *Econ. Geol.* **69**, 293–317 (1974).
- Nesbitt, H. W. & Young, G. M. *Nature* **299**, 715–717 (1982).
- Kimberley, M. M. *MAC Short Course in Uranium Deposits* (ed. Kimberley, M. M.) 339–381 (Mineral Association of Canada, 1978).
- Dimroth, E. *Phil. Trans. R. Soc. A* **291**, 277–287 (1979).
- Robinson, A. G. & Spooner, E. T. *Geol. Soc. Am. Abstr. Prog.* **13**, 539 (1981).
- Thode, H. G., Dunford, H. B. & Shima, M. *Econ. Geol.* **57**, 565–578 (1962).
- Hoefs, J., Nielsen, H. & Schildowski, M. *Econ. Geol.* **63**, 975–977 (1968).
- Chukhrov, F. V., Vinogradov, V. I. & Ermilova, L. P. *Miner. Depos.* **5**, 209–222 (1970).
- Pretorius, D. A. *Handbook of Strata-bound and Stratiform Ore Deposits* Vol. 7 (ed. Wolf, K. H.) 29–87 (Elsevier, Amsterdam, 1976).
- Smith, N. D. & Minter, W. E. L. *Econ. Geol.* **75**, 1–14 (1980).
- Utter, T. *Geol. Rdsch.* **67**, 774–805 (1978).
- Hallbauer, D. K., Jahns, H. M. & Beltmann, H. A. *Geol. Rdsch.* **66**, 477–491 (1977).
- Hallbauer, D. K. *Geol. Surv. Prof. Paper* **1161-M** (1981).
- Young, G. M. *Can. J. Earth Sci.* **12**, 1250–1254 (1975).

Peroxyacetyl nitrate in the free troposphere

Hanwant B. Singh & Louis J. Salas

Atmospheric Science Center, SRI International, Menlo Park, California 94025, USA

First free tropospheric (0–8 km) peroxyacetyl nitrate ($\text{CH}_3\text{C}(\text{O})\text{OONO}_2$, PAN) measurements in the clean marine air of the Pacific (36–40°N) show that the tropospheric reservoir of nitrogen oxides (NO , NO_2) may be 30–50% larger than hitherto believed. We show here that these measurements provide experimental support for the contention that PAN is an important constituent of the natural atmosphere, and a potentially significant reservoir of reactive nitrogen. There is evidence to suggest that higher PAN concentrations exist in the free troposphere over land, as compared with the marine environment. Preliminary indications are that measured PAN was produced through *in situ* photochemical mechanisms, and was not transported from polluted sources.

PAN was first detected and identified by Stephens and co-workers^{1,2} during studies designed to understand better the processes of smog formation. In recent years, the formation, transport and chemistry of this phytotoxin have been studied extensively^{3–8}. Laboratory studies of PAN chemistry have demonstrated that PAN exists in equilibrium with NO_2 [$\text{CH}_3\text{C}(\text{O})\text{OONO}_2 \rightleftharpoons \text{CH}_3\text{C}(\text{O})\text{OO} + \text{NO}_2$], and its thermal

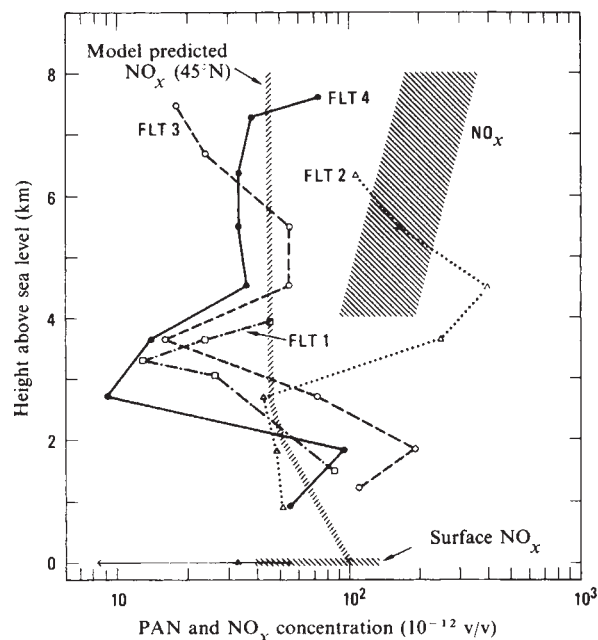


Fig. 1 Distribution of PAN in the free troposphere. Data points joined by lines show PAN concentrations. Flight positions are between 36.8° and 39.5° N latitude and 123° to 125° W longitude, approximately 60 to 150 km west of the California coast. Flights 1, 2 and 4 are daytime flights (1200–1800 h) on 8, 10, and 14 September, respectively. Flight 3 is a night-time flight (13–14 September) with all data collected during 2200–0200 h. Surface PAN data are averages of those presented in Fig. 2. The shaded area represents the range of mid-latitude NO_x concentrations from other studies^{11,12}. The model predicted NO_x profile is the zonal average at 45° N (ref. 18).

decomposition is highly temperature dependent^{4,5}. Light hydrocarbons and reactive nitrogen oxides (NO_x) have also been measured as ubiquitous components of the global atmosphere^{9–12}. Based on existing knowledge of tropospheric photochemistry, it has been proposed that naturally occurring non-methane hydrocarbons (NMHCs) and NO_x could lead to *in situ* synthesis of significant PAN concentrations, especially in the colder regions of the middle and upper troposphere^{13,14}. At these predicted concentrations, PAN would provide an important reservoir for reactive nitrogen oxides (and of peroxy radicals) which are thought to have a central role in the chemistry of the troposphere⁸.

The distribution of PAN in the unpolluted atmosphere is not known. Virtually all measurements to date have been conducted in the polluted air. In our laboratory, a methodology was developed for quantitative PAN analysis to a mixing ratio of 5×10^{-12} v/v. A sensitivity of 1×10^{-12} v/v was easily achieved. Details involving sampling, calibration and detection are given elsewhere¹⁵.

Briefly, PAN was measured with the help of an electron capture gas chromatograph, designed for both ground and airborne operation. A twin engine aircraft (Beechcraft B-80/8800 Queen Air—ceiling altitude of 8 km) was used for the airborne studies. A 50–200 ml (1 atm, 20 °C) volume of air was preconcentrated in an unpacked 0.15 cm o.d. stainless tube of 1.24 ml volume, at liquid argon temperature before analysis. Separation was achieved on a 36 cm $\times$ 3 mm i.d. Teflon column packed with 10% carbowax 600 on 80/100 mesh Supelcoport. A column temperature of 33 °C, detector temperature of 64 °C, and a carrier gas (5% Ar/CH_4) flow rate of 30 ml min⁻¹ were used. Ambient PAN identity was confirmed by comparing several properties with PAN synthesized in the laboratory. These were (1) retention times on two gas-chromatography columns, (2) electron ionization efficiency determinations, (3) variation of PAN response with changing electron capture detector (ECD) temperatures, (4) PAN destruction when passed through a 1% KOH scrubber, and (5) studies of PAN

loss when passed through heated metal tubes. In all these cases, ambient PAN behaved identically to synthesized PAN. Analysis of blanks ensured the absence of any artefact problems.

An all-glass continuous PAN calibration system based on the UV photolysis (black light, λ_{max} 360 nm) of $\text{CH}_3\text{CHO}/\text{Cl}_2/\text{NO}_2$ (ref. 16) was devised. The trimixture was irradiated as it flowed through a glass reactor with a typical residence time of 4 s. The PAN yield was determined by measuring acetaldehyde loss. Depending on the measured NO/NO_2 ratio at the reactor exit, the system was 90–95% efficient. Acetaldehyde was measured by a gas chromatographic-flame ionization detection technique with a $\pm 0.7\%$ precision. PAN produced in this fashion at $300\text{--}400 \times 10^{-9}$ v/v levels was further diluted in a dynamic two-stage all-glass dilution system. PAN produced by the reactor was also independently measured by placing the reactor in line with a PAN scrubber and a chemiluminescent NO_x detector. The PAN scrubber utilized glass beads coated with 1% KOH solution, and could remove PAN with an efficiency of 99–100%. PAN was determined from the post scrubber NO_x loss when the reactor lights were turned on. These independent checks led to agreement within 10% (ref. 15). Measurement precision was $\pm 3\%$, and an overall accuracy of $\pm 20\%$ is estimated. PAN calibrations were also performed in a pressure chamber to correct for the altitude dependence of the instrument response¹⁵.

All PAN data were collected in real time. Figures 1 and 2 present the results of PAN measurements to an altitude of 8 km over the Pacific Ocean. The surface PAN data were obtained during 25–29 August 1982, at a Pacific coastal site located at the lighthouse station at Point Arena, California ($38^\circ 57' \text{N}$, $123^\circ 44' \text{W}$). Aircraft measurements were conducted $\sim 60\text{--}150$ km west of the California coast ($36.8\text{--}39.5^\circ \text{N}$) during September 1982 (Fig. 1). In all cases, the objective was to sample the cleanest possible marine air. Superimposed on Fig. 1 are the ranges of measured NO_x concentrations over midlatitudes which have been presented recently¹¹. Because no reliable midlatitude surface level Pacific NO_x data are available, the diurnal behaviour of PAN (Fig. 2) is compared with the results of a similarly-situated Atlantic marine site on the west coast of Ireland¹². Measurements of light hydrocarbons over the Atlantic⁹ and the Pacific¹⁰ suggest comparable precursor abundance.

It is clear from Fig. 1 that compared with NO_x , PAN is present in significant concentrations at all levels up to 8 km. Although both PAN and NO_x show substantial variability, the abundance of PAN can be easily 30–50% of NO_x . The true PAN abundance relative to NO_x is probably higher, because many chemiluminescent NO_x instruments suffer from a positive PAN interference^{12,17}. Indeed, the measured PAN levels are comparable with the model-predicted NO_x levels at midlatitudes¹⁸. The data, however, are too sparse to allow a general indication of its vertical distribution. The mean surface PAN concentration (in units of 10^{-12} v/v) of 32 ± 24 is an important fraction of the mean NO_x concentration of 87 ± 47 (Fig. 2). Because of interference from species such as PAN and nitric acid¹², the true NO_x abundance is significantly less than 87 parts in 10^{12} parts of air.

The surface PAN shows a distinct diurnal variation that is similar to that of NO_x (Fig. 2). The reasons for the significant NO_x diurnal variation are not fully understood¹². PAN is slowly destroyed in contact with surfaces¹⁹, causing the night-time concentrations to be lower. Hydrocarbon measurements at the Point Arena site show the presence of 0.1–0.2 parts in 10^9 parts of air of ethene and propene, and somewhat higher levels of the less reactive $\text{C}_2\text{--C}_5$ alkanes¹⁰. The alkenes are globally distributed^{9,10} and probably of oceanic origin²⁰. Thus, some PAN could be synthesized within hours by reactions involving highly reactive alkenes and nitrogen oxides in a manner similar to polluted atmospheres². Alkanes would also participate in these processes but at slower rate¹³. The results of the night-time aircraft flight (Fig. 1) indicated that PAN is preserved aloft. The night-time PAN levels (flight 3) were higher than the

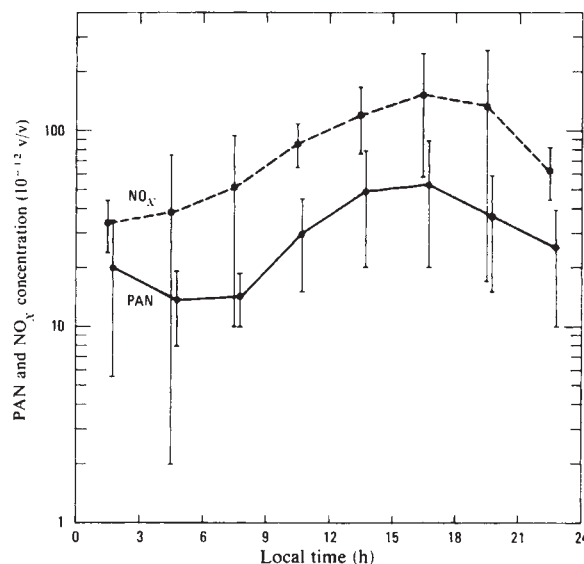


Fig. 2 Surface level diurnal behaviour of PAN in a marine air environment. PAN data were collected at a mid-latitude coastal Pacific site at Point Arena, California ($38^\circ 57' \text{N}$; $123^\circ 44' \text{W}$) during 25–29 August 1982. The NO_x data are from a similarly situated North Atlantic Coastal site on the west coast of Ireland during 13–16 June and 20–23 June 1979¹². Both PAN and NO_x concentrations are 3-h averages ($\pm 1\sigma$). The PAN data are shifted slightly in time for clarity.

daytime levels of the following day, except above 6 km. At present, data do not allow the inference of a PAN diurnal pattern in the free troposphere.

Note that the surface PAN concentrations reflect weather conditions of overcast stratus, with occasional fog and drizzle. The solar flux at the experiment site never exceeded $1 \text{ cal cm}^{-2} \text{ min}^{-1}$. It is possible, therefore, that higher PAN levels may be found at the surface on days with greater solar flux availability. Infrequently, peroxypropionyl nitrate could also be detected at concentrations that did not exceed 7% of PAN.

Two sources of PAN are possible: *in situ* photochemical formation, or long distance transport from polluted sources^{8,13,14}. It also has been suggested that heterogeneous processes could lead to PAN synthesis, but the atmospheric significance of such processes is not known²¹. The present data when interpreted in the context of prevailing meteorology support the contention that the bulk of the PAN was not transported from pollution sources. First, the locations for aircraft and ground experiments were those that were strongly influenced by Pacific westerly winds. During the Point Arena experiment, winds were west to north-west ($255^\circ\text{--}345^\circ$) about 90% of the time, and northerly the remaining 10%. LFM II (Limited area Finemesh Model) 24-h back trajectories (twice daily) for the surface, 850, and 700 mbar levels from NOAA, further confirmed the marine origin of these air masses. When all data during periods of northerly winds were excluded, PAN concentrations changed negligibly. In addition, O_3 , CCl_3F and NO_x were measured at this site. The first two were at their expected background concentrations ($\text{O}_3 = 34 \pm 18$ parts in 10^9 ; $\text{CCl}_3\text{F} = 200 \pm 10$ parts in 10^{12}), and NO_x could not be detected (<1 part in 10^9) by our instrument. Trajectories for the aircraft experiments are not possible because of lack of data over the Pacific. However, meteorological soundings from a nearby land station (Oakland, California) indicated that winds were always from the south-west to the north-west region ($200^\circ\text{--}330^\circ$) at all altitudes during flights 1, 3 and 4. Flight 2 wind directions were northerly (wind direction of 20° at 1 km and $345\text{--}360^\circ$ aloft), and land contact was possible. In all of these cases, synoptic weather patterns precluded any direct possibility of PAN transport from a polluted city. It is possible, however, that land air masses because of a higher burden of reactive hydrocarbons and NO_x compared with the marine environment²², may be able to synthesize larger PAN concentrations²³.

In May 1978 at a remote continental site at Jetmore, Kansas (lat. 37°54' N; long. 99°53' W), an average PAN concentration of 253 (± 103) parts in 10^{12} parts of air was measured²³. During Flight 2, PAN concentrations measured at 6.4 km altitude 90 km west of the San Francisco coastal line (37° N) were roughly twice those measured at a 175 km west location at the same altitude.

The most probable source of measured tropospheric PAN is photochemistry involving reactions of non-methane hydrocarbons and NO_x (refs 13, 14). The oceanic origin of alkenes, and global presence of other hydrocarbons, lead us to suggest that PAN is most likely a natural constituent of the global atmosphere and an important reservoir for nitrogen oxides. The existence of PAN in the free troposphere provides a qualitative validation of the photochemical theories. Additional measurements are required to elucidate further the distribution of PAN in the global atmosphere, and its role as a reservoir of reactive nitrogen and of free peroxy radicals. As PAN can be measured with such sensitivity, and is a unique product of complex photochemistry, it can have a special role in the future validation of global photochemical models.

Research was funded by NSF under grant ATM-8111758.

Received 17 November 1982; accepted 20 January 1983.

- Stephens, E. R., Hanst, P. L., Doerr, R. C. & Scott, W. E. *Ind. Engng Chem.* **48**, 1498–1504 (1956).
- Stephens, E. G. *Adv. env. Sci.* **1**, 119–146 (1969).
- Nieboer, H. & van Ham, J. *Atmos. Envir.* **10**, 115–120 (1976).
- Hendry, D. G. & Kenley, R. A. *J. Am. chem. Soc.* **99**, 3198–3199 (1977).
- Cox, R. A. & Roffey, M. J. *Envir. Sci. Technol.* **11**, 900–906 (1977).
- Nielsen, T., Samuelsson, U., Grennfelt, P. & Thompson, E. *Nature* **293**, 553–555 (1981).
- Lonneman, W. A., Bufalini, J. J. & Seila, R. L. *Envir. Sci. Technol.* **10**, 374–380 (1976).
- Crutzen, P. J. A. *Rev. Earth planet. Sci.* **7**, 443–472 (1979).
- Rudolph, J., Ehhalt, D. H. & Tonnisson, A. *J. geophys. Res.* **86**, 11959–11964 (1981).
- Singh, H. B. & Salas, L. J. *Geophys. Res. Lett.* **9**, 842–845 (1982).
- Kley, D., Drummon, J. W., McFarland, M. & Liu, S. C. *J. geophys. Res.* **86**, 3153–3161 (1981).
- Helas, G. & Warneck, P. *J. geophys. Res.* **86**, 7283–7290 (1981).
- Singh, H. B. & Hanst, P. L. *Geophys. Res. Lett.* **8**, 899–903 (1981).
- Aikin, A. C., Herman, J. R., Maier, E. J. & McQuillan, C. J. *J. geophys. Res.* **87**, 3105–3118 (1982).
- Singh, H. B. & Salas, L. J. *Atmos. Envir.* (in the press).
- Gay, B. W., Noonan, R. C., Bufalini, J. J. & Hanst, P. L. *Envir. Sci. Technol.* **10**, 83–85 (1976).
- Winer, A. M., Peters, J. W., Smith, J. P. & Pitts, J. N. *Envir. Sci. Technol.* **8**, 1118–1122 (1974).
- Logan, J. A., McElroy, M. B., Wofsy, S. C. & Prather, M. J. *J. geophys. Res.* **96**, 7210–7254 (1981).
- Garland, J. A. & Penkett, S. A. *Atmos. Envir.* **10**, 1127–1131 (1976).
- Swinerton, J. W. & Lamontagne, R. A. *Envir. Sci. Technol.* **8**, 657–663 (1974).
- Lovelock, J. E. & Penkett, S. A. *Nature* **249**, 434 (1974).
- Robinson, E. *Pure appl. Geophys.* **116**, 372–384 (1978).
- Singh, H. B. *et al.* EPA-600/3-79-107 (1979).

Marine biological controls on atmospheric CO₂ and climate

M. B. McElroy

Center for Earth and Planetary Physics, Harvard University, Cambridge, Massachusetts 02138, USA

It is argued here that the concentration of fixed nitrogen in the ocean is not in steady state, that the ocean gains N during times of glacial advance, losing it when glaciers recede. The abundance of atmospheric CO₂ responds to changes in marine biological productivity, and it is proposed that climatically significant fluctuations in CO₂ on time scales of 10³–10⁴ yr are controlled mainly by variations in oceanic N.

Carbon dioxide has an important influence on the radiative budget of the atmosphere¹. Its concentration is variable: it is increasing currently² by about 1 p.p.m. yr⁻¹ due to burning of fossil fuel and changing patterns of land use. Studies of ice cores indicate that CO₂ fluctuated naturally in the past by as much as a factor of 2 on time scales as brief as a few thousand years^{3–5}. The concentration was low during the last ice age, raising obvious questions concerning the role of the gas as a causative factor in climate change. Change in CO₂ on time scales less than 10⁵ yr must reflect an adjustment in the distribution of carbon between the ocean, atmosphere and biosphere⁶.

Incorporation of carbon into marine sediments, and other geological processes, require more than 10⁵ yr to change significantly the abundance of CO₂ in the atmosphere–ocean–biosphere system.

Broecker⁶ points out that deep ocean water, transported isochemically to the surface, would be expected to release significant quantities of CO₂ as it warms up to ambient temperature—the equilibrium concentration of CO₂ could be as high as 1,000 p.p.m. The lack of equilibrium between the upper ocean–atmosphere–biosphere on the one hand and the lower ocean on the other reflects in large part the importance of biological activity. Photosynthesis in near-surface waters supplies a rain of detrital carbon which passes through the thermocline, representing non-conservative transport of carbon from the upper to lower ocean.

More than 80% of carbon fixed annually by marine biota is fixed in the open ocean^{7,8}. The productivity of coastal areas is ~2–4 times that of the open ocean, where fixation averages typically 50 g C m⁻² yr⁻¹. Productivity in upwelling regions may reach as high as 300 g C m⁻² yr⁻¹ with values as large as 1,000–2,000 g cm⁻² yr⁻¹ observed on occasions off Peru⁹. Regions of intense upwelling account, however, for only 0.1% of the ocean surface; coastal areas occupy as much as 10%. Productivity is limited by the supply of nutrients, with attention focused usually on the macronutrients N and P. There is considerable dispute as to which of these elements is limiting: biologists^{8,10} favour N; geochemists⁶, P. Both sides may be correct. It appears that nitrogen is the limiting nutrient for life in the ocean today. Phosphorus may play a critical role on time scales of 10⁵ yr and longer, however.

Phosphorus is derived ultimately from weathering of continental rocks. It accounts for about 0.1% of this material¹¹. It follows that the flux to the ocean should be ~2 × 10¹³ g yr⁻¹. Only a small fraction of this P, about 5%, is available to the marine biota¹². The balance passes directly into sediments. The flux of biologically useful 'reactive' P, supplying the oceanic reservoir of dissolved PO₄³⁻–P (8 × 10¹⁶ g), is about 10¹² g P yr⁻¹, indicating a residence time for oceanic PO₄³⁻–P of about 10⁵ yr, similar to the value for the carbon content of the ocean plus atmosphere plus biosphere.

There are three potential sources of fixed oceanic N: transport from the land in rivers and estuaries; input from the atmosphere, primarily through rain; and *in situ* fixation. Nitrogen is lost mainly by denitrification with additional removal to sediments.

Denitrification is confined to a relatively small part of the world's oceans where demand for O₂ results in almost complete anoxia. Waters in the Eastern Tropical Pacific are particularly important. Several methods have been applied to estimate loss of fixed nitrogen from anaerobic waters. One makes use of the concept of missing nitrogen. The advective history of a water mass is inferred from measurements of the spatial distribution of quasi-passive tracers such as salinity. A reduction in the concentration of dissolved N is observed for waters in which O₂ drops below about 0.1 mg l⁻¹. The method implies a loss of N equivalent to about 2.5 × 10¹³ g N yr⁻¹ from the Eastern Tropical South Pacific and similar results are obtained from *in situ* studies of rates for microbial metabolism¹³. Low O₂ waters of the Eastern Tropical North Pacific account for removal of about 3 × 10¹³ g N yr⁻¹. Denitrification in the Eastern Pacific is more important than denitrification in the Arabian Sea¹⁴ and is larger than the loss from pelagic sediments¹⁵. Nitrogen is lost to sediments at a rate of about 2 × 10¹³ g N yr⁻¹ (ref. 16).

It follows that the lifetime of oceanic N is a little less than 10⁴ yr. One is tempted to argue that variations in climate on time scales of order 10⁴ yr are modulated by variations in CO₂ responding to fluctuations in oceanic N. Broecker⁶ argued that these variations in CO₂ are controlled for the most part by changes in P, attributing the short time constant (10⁴ yr) to exchange of P with coastal sediments. It must be pointed out that episodic uptake and release of nutrients by coastal sediments should involve N as well as P. The 10⁴ yr time constant arises naturally for N, but is *ad hoc* for P. We believe that the

evidence favours the conclusion that N is the limiting nutrient for marine biota on time scales of order 10^4 yr and less. The situation is less clear on longer time scales where P may assume larger importance.

Loss of oceanic N to sediments and to the atmosphere requires a compensating source, with production equalling loss, at least on time scales of the order of 10^4 yr. We shall argue that the contemporary source is too small, that the ocean today is losing N at a rate faster than it is supplied, and that the imbalance is reversed during glacial times when sea level is lower. Supply from land could be more efficient then; erosion of coastal sediments would provide additional input of stored nutrients; and *in situ* fixation could be more important.

Fixed N is formed in the atmosphere by lightning. Our understanding of atmospheric chemistry has developed rapidly over the past decade and we can be confident that the source of marine NO_3^- from the atmosphere is small, less than 2×10^{13} g N yr $^{-1}$, perhaps as little as 2×10^{12} g N yr $^{-1}$. There may be some transfer of N from land to ocean through the atmosphere as NH_3 . Available data suggest that this also is small. We adopt a value of 8×10^{12} g yr $^{-1}$ for the flux of N from air to sea at the present time. Precipitation can account for no more than 10% of the time averaged source of oceanic N.

The literature contains a variety of estimates for fixation of N in the ocean. Attention has focused recently on the relatively widespread planktonic blue-green alga *Oscillatoria trichodesmium*¹⁷⁻²². *Oscillatoria* grows slowly; a doubling time of 180 days is reported by McCarthy and Carpenter²². There is evidence that growth is limited at ambient levels of oceanic PO_4^{3-} by availability of P. *Oscillatoria* can provide a significant source of N in localized regions at some times—the Eastern Caribbean, for example²⁰. Its importance on a global scale is limited, however. We agree with Fogg's²³ conclusion that biological fixation in the ocean can account for no more than 1.5×10^{13} g N yr $^{-1}$. The value 5×10^{12} g N yr $^{-1}$ derived by Capone and Carpenter²⁴, based on a survey of all of the available data for *Oscillatoria*, is adopted for present purposes. It follows that the combined input from rain and fixation can account for no more than 15% of the time averaged oceanic requirement for N.

The flux of N in rivers could be as large as 4×10^{13} g N yr $^{-1}$, but is more probably half this, about 2×10^{13} g N yr $^{-1}$ (ref. 25 and M.B.M. and F. Knox, in preparation). We arrive at these numbers as follows: data for the 20 largest rivers indicate a flux of NO_3^- -N of about 6×10^{12} g N yr $^{-1}$. These rivers account for approximately 36% of the net flow of water to the ocean, suggesting a global flux of NO_3^- -N of 1.6×10^{13} g N yr $^{-1}$. Dissolved organic N could contribute an additional 8×10^{12} g N yr $^{-1}$, for a total of 2.4×10^{13} g N yr $^{-1}$. Rivers could contribute between 20% and 40% of the ocean's average requirements for N. A rather small fraction of nitrogen in today's rivers is transferred directly to the ocean, however, and transfer of N fixed in coastal marshes²⁴ should be similarly difficult. Most of the nitrogen in rivers today is trapped in deltas and coastal sediments^{26,27} but could be transferred to the open ocean when sea level is lower and erosion more efficient⁶.

One can speculate now on the course of the oceanic N budget over a typical 10^4 yr climatic oscillation. Advance of the ice sheet during the cooling phase will be accompanied by increased erosion, with transfer to the ocean of terrestrial N produced in part by enhanced fixation in moraine environments near the ice margin. It would be associated also with lowering of sea level, erosion of coastal sediments with consequent addition of nutrients to the deep ocean, promoting greater biological productivity, leading to reduction in CO_2 , enhancing climatic cooling—positive feedback. Supply of N to the ocean would dominate loss during this epoch. Supply would diminish as sediments were depleted and as sea level stabilized with migration of the ice sheet to lower latitudes. Loss of oceanic N would assume increasing importance. The ocean would become less productive, returning CO_2 to the atmosphere, promoting climatic warming with positive feedback on the warming as on

the cooling phase of the climatic cycle. There are several factors which might combine to effect the transition from warming to cooling. In the first place, the 10^4 yr time constant emphasized here for N was derived empirically. Biological productivity could decrease during the warming phase and conditions more suitable for elevated growth of *Oscillatoria*, high levels of PO_4^{3-} in surface waters, for example, could lead to enhanced fixation of nitrogen. Ice surges could cause abrupt changes in sea level with consequent mobilization of nutrients in fertile coastal regions or the transition could arise for purely physical reasons associated, for example, with saturation of the coastal zone's ability to store sediment.

Thus denitrification in upwelling systems represents a significant sink for oceanic fixed N on time scales of 10^3 – 10^4 yr. There is no obvious reason to expect an instantaneous steady state and the available evidence suggests that the ocean is losing N now at a rate faster than it can be supplied. Variations in productivity associated with changes in nitrogen are expected to influence the distribution of carbon between the atmosphere–biosphere–upper ocean and the lower ocean. Episodic changes in oceanic N can account for the variability of CO_2 inferred from analyses of ice cores and may be expected to have a significant influence on climate.

This work was supported by NSF grant DEB79-20282 and NASA grant NAG1-55 to Harvard University. I thank G. E. Fogg, H. D. Holland, F. Knox, J. J. McCarthy, J. J. Walsh, and S. C. Wofsy for useful discussions.

Received 19 August 1982; accepted 10 January 1983.

1. Hansen, J. *et al.* *Science* **213**, 957–966 (1981).
2. Keeling, C. D. & Bacastow, R. B. *Energy and Climate*, 72–95 (National Academy of Sciences, Washington, D.C., 1977).
3. Delmas, R. J., Ascencio, J. M. & Legrand, M. *Nature* **284**, 155–157 (1980).
4. Berner, W., Oeschger, H. & Stauffer, B. *Radiocarbon* **22**, 227–235 (1980).
5. Neftel, A., Oeschger, H., Schwander, J., Stauffer, B. & Zumbund, R. *Nature* **295**, 220–223 (1982).
6. Broecker, W. S. *Prog. in Oceanogr.* **11**, 151–197 (1982).
7. Ryther, J. H. *Science* **166**, 72–76 (1969).
8. McCarthy, J. J. (in preparation).
9. Walsh, J. J. *Nature* **290**, 300–304 (1981).
10. Perry, M. J. & Eppley, R. W. *Deep Sea Res.* **28A**, 39–49 (1981).
11. Mason, B. *Principles of Geochemistry* 3rd edn (Wiley, New York, 1966).
12. Froelich, P. N., Bender, M. L., Luedtke, N. A., Heath, G. R. & DeVries, T. *Am. J. Sci.* **282**, 474–511 (1982).
13. Codispoti, L. A. & Packard, T. T. *J. mar. Res.* **38**, 453 (1980).
14. Deuser, W. G., Ross, E. H. & Modzinska, Z. I. *Deep Sea Res.* **25**, 431–445 (1978).
15. Bender, M. L., Fanning, K. A., Froelich, T. N. & Heath, G. R. *Science* **198**, 605 (1977).
16. Holland, H. D. *The Chemistry of the Atmosphere and Oceans* (Wiley, New York, 1978).
17. Dugdale, R. C., Goering, J. J. & Ryther, J. H. *Limnol. Oceanogr.* **9**, 507–510 (1964).
18. Goering, J. J., Dugdale, R. C. & Menzel, D. W. *Limnol. Oceanogr.* **11**, 614–620 (1966).
19. Carpenter, E. J. & McCarthy, J. J. *Limnol. Oceanogr.* **20**, 389–401 (1975).
20. Carpenter, E. J. & Price, C. C. *Limnol. Oceanogr.* **22**, 60–72 (1977).
21. Fogg, G. E. *Ecol. Bull. Stockholm* **26**, 11–19 (1978).
22. McCarthy, J. J. & Carpenter, E. J. *J. Phycol.* **15**, 75–82 (1979).
23. Fogg, G. E. *Phil. Trans. R. Soc. B* **296**, 511–520 (1982).
24. Capone, D. & Carpenter, E. J. *Science* **217**, 1140 (1982).
25. Meybeck, M. *Am. J. Sci.* **282**, 401–450 (1982).
26. Edmond, J. M., Boyle, E. A., Grant, B. & Stallard, R. F. *Deep Sea Res.* **28A**, 1339–1374 (1981).
27. Gibbs, R. J. *Geology* **9**, 77–80 (1981).

Particle fluorescence and bioluminescence distributions in the eastern tropical Pacific

William W. Broenkow, Alan J. Lewitus, Mark A. Yarbrough & Robert T. Krenz

Moss Landing Marine Laboratories, Moss Landing, California 95039, USA

The measurement of certain optical properties provides valuable information regarding the distribution of planktonic organisms in the ocean. Here we report vertical profiles of fluorescence, beam attenuation and bioluminescence in the oxygen minimum zone of the eastern tropical Pacific relative to possible physical and biological particle sources. Of particular interest is the observation of a deep fluorescence maximum within the oxygen minimum zone that has not been previously reported.

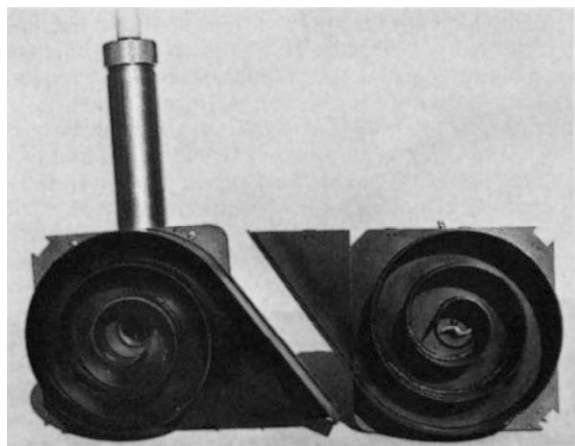


Fig. 1 Bioluminescence sensor and light baffle. Baffle has been opened to show window to photomultiplier housing (left) and impeller (right).

Data were taken with a conductivity, temperature, dissolved oxygen, depth (CTOD) profiler with an auxiliary fifth channel, which was time-shared on different casts by a Variosens fluorometer¹, a beam transmissometer, and a bioluminescence sensor designed by us. Bioluminescence was measured within a light baffle by a photomultiplier tube chosen for a spectral response that closely matches the luminescence characteristics of most marine organisms. Our spiral design of the light baffle (Fig. 1) is unique, we believe, because it completely excludes ambient light, yet is short enough to give a 4-s residence time within the chamber, and it produces nearly turbulence-free flow which avoids pre-stimulation of potentially luminescent organisms before they reach the photomultiplier chamber. Flow is accelerated through the baffle by a funnel, and bioluminescence is stimulated by an impeller located directly in front of the photomultiplier window.

These instruments allowed us to investigate water column distributions of different particle components. The fluorometer is routinely used to assess phytoplankton biomass distributions from *in vivo* chlorophyll *a* ^{2,3} fluorescence. Without extensive calibration, this procedure may not produce accurate results, because *in vivo* and *in situ* chlorophyll *a* quantum efficiency is highly variable⁴⁻⁶, and one cannot always discount the importance of fluorescence by dissolved organic materials⁷. The total attenuation coefficient, *c*, measured by the beam transmissometer is related to the concentration of suspended and dissolved materials⁸, and in the open ocean, *c* has been shown to correlate with phytoplankton standing stock⁹, and with chlorophyll *a* fluorescence¹⁰. Bioluminescence measurements are, perhaps, the most specific bio-optical survey tool, because even though a wide variety of organisms (bacteria, dinoflagellates, some zooplankton and fish) luminesce¹¹, inorganic particulate and dissolved materials are not detected.

Our observations in the oxygen minimum region off central Mexico (Fig. 2) showed an interesting juxtaposition of bio-optical distributions. The dominant fluorescence feature was the primary maximum at 60 m (range 15–110 m) which was strongly correlated with chlorophyll *a* concentrations measured by extracted *in vitro* methods¹¹. Maximum chlorophyll *a* concentrations of 0.6 mg m^{-3} occurred at the fluorescence maximum, which was near the 0.1% relative irradiance level in the centre of the main thermocline (Fig. 2a, b). A second fluorescence maximum was observed near 130 m in the oxygen minimum at 12 of the 20 profiles obtained (Fig. 2b), and a third broad fluorescence feature between 200 and 400 m was found at 17 stations (Fig. 2c). A transmissometer profile (Fig. 2c) showed a close correlation with the fluorescence profile taken at the same station a day earlier. The primary fluorescence

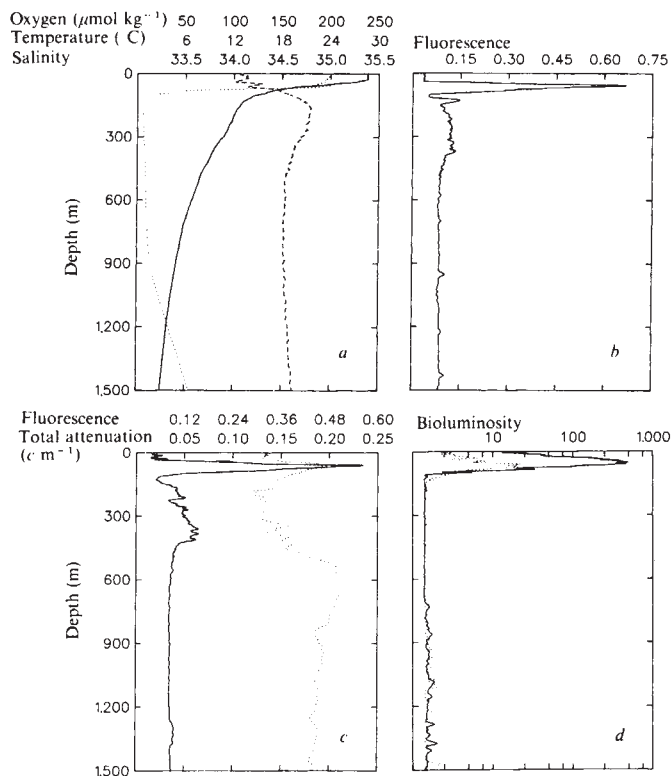


Fig. 2 Typical vertical distributions: *a*, temperature (solid line), salinity (dashed line) and dissolved oxygen ($\mu\text{mol kg}^{-1}$) 26 October, $18^\circ \text{N } 108^\circ \text{W}$; *b*, fluorescence 9 November, $17^\circ 46' \text{N } 109^\circ \text{W}$; *c*, fluorescence (solid line) 27 October and total attenuation coefficient (dotted line) 26 October, $18^\circ \text{N } 108^\circ \text{W}$; *d*, bioluminescent intensity daytime (dotted line) 6 November, $17^\circ 35' \text{N } 108^\circ 54' \text{W}$ and night-time (solid line) 7 November, $17^\circ 35' \text{N } 108^\circ 27' \text{W}$.

maximum coincided with an attenuation maximum. Between 180 and 420 m, the total attenuation minimum was associated with the deep fluorescence maximum (Fig. 2c). Within this layer, *c* was directly correlated with fluorescence ($r = 0.65$). The bioluminescence profile (Fig. 2d) showed a strong maximum at the primary fluorescence maximum, relatively little activity in the oxygen minimum zone between 120 and 700 m (where oxygen concentrations were $< 20 \mu\text{mol kg}^{-1}$), and sporadic but increasing activity below 700 m. The near-surface bioluminescence distribution showed strong diurnal periodicity with larger night-time intensities throughout the photic zone (Fig. 2d).

These observations result from the complex interplay of physics and biology, and cannot be fully explained without further study, but some conclusions are in order.

The nocturnal increase in bioluminescence activity has been attributed to an endogenous rhythm in luminescent dinoflagellates¹², and may also be caused by the stimulation of dinoflagellates by zooplankton¹³ ascending at night. A sound-scattering layer was observed at the study site to migrate vertically from 140 m (well within the oxygen minimum) at sunset to 50 m 1 h after sunset (Fig. 3a). Bioluminescence time series were made at 50 and 100 m (Fig. 3b) at depths of high and low luminescent activity, respectively.

These data showed that the passage of the sonic scattering layer did not result in increased bioluminescence at 100 m, but that a marked increase in bioluminescence occurred simultaneously with the passage of the scattering layer past the 50-m level. This suggests that the increased bioluminescence was caused by stimulation of organisms already present at that depth, and that we did not detect luminescence from the unidentified organisms comprising the scattering layer.

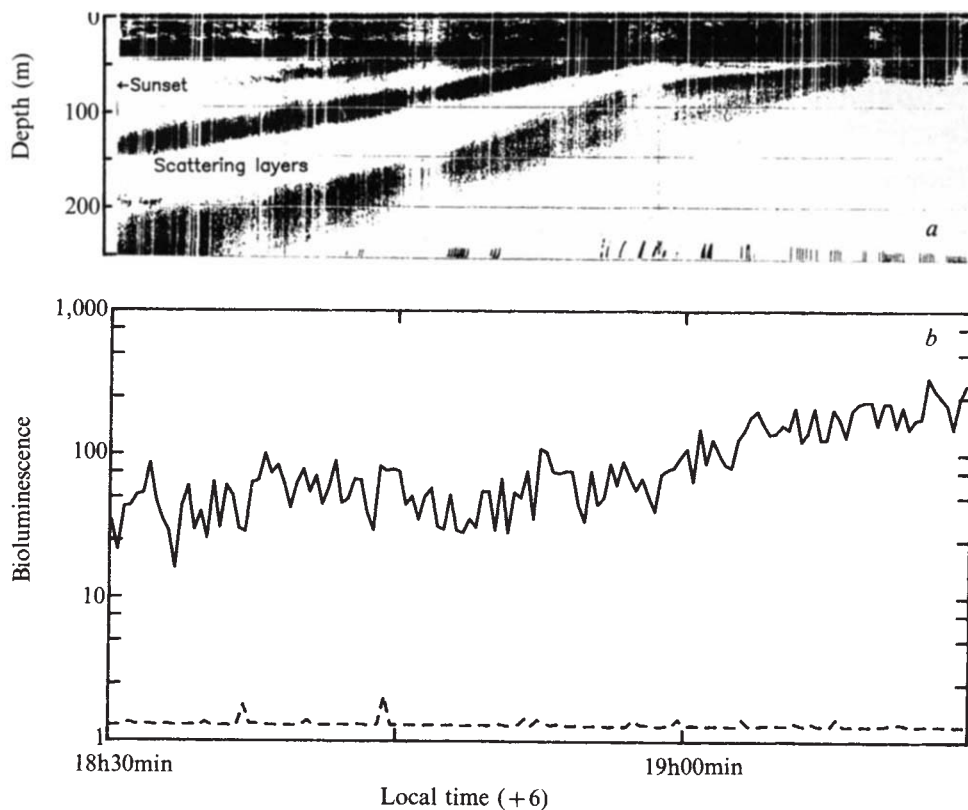


Fig. 3 *a*, Fathometer trace 7 November, 17°55' N 108°49' W beginning at sunset; *b*, bioluminescence time series at 50 and 100 m. The 50-m time series (solid line) was obtained simultaneously with the fathometer trace, and the 100-m data (dashed line) were obtained 2 days later. Data were digitized at 2-s intervals and have been averaged to 20-s intervals for plotting.

Dinoflagellates (*Ceratium* spp.) were relatively abundant in the euphotic zone (M. D. Tuel, personal communication). The lack of bioluminescent activity in the oxygen minimum agrees with the lack of net plankton obtained from this zone (G. A. Knauer and M. D. Tuel, personal communication), and qualitatively with the low total attenuation coefficient observed there.

The primary fluorescence maximum at 60 m contrasted with the low surface fluorescence (Fig. 2*b, c*). The lowest fluorescence values in the entire water column were observed at the surface, and reflected both the low phytoplankton standing stocks and strong photo-inhibition^{5,14}.

The secondary fluorescence maximum (at ~130 m) occurred almost precisely at the depth of minimum water-column oxygen concentration. Because it was found predominantly along the eastern and southernmost sides of the study area, we believe the feature was probably caused by advection. Waters in the oxygen minimum were not stagnant. A cyclonic eddy dominated geostrophic flow at the study site. Speeds to 20 cm s⁻¹ and strong horizontal and vertical shear were observed in the oxygen minimum at 120 m (ref. 15).

The deep fluorescence maximum between 200 and 400 m was an unexpected and dramatic feature apparently associated with the oxygen minimum zone. Fluorescence levels in the deep maximum were three to four times those observed at the surface. Several hypotheses may be formed to account for this feature, which we believe has never been observed previously: (1) its association with the oxygen minimum zone suggests that the fluorescence may be caused by dissolved or particulate organic material preserved in oxygen-impoverished waters. (2) The fluorescence maximum may be due to undigested faecal plant pigments originating from overlying surface waters¹⁶. (3) This layer may represent fluorescent pigments advected from source regions remote from the study site. The observation that fluorescence and *c* covaried and showed similar microstructure (Fig. 2*c*) suggests that the fluorescence was associated with particles rather than dissolved organic material.

We find it more than coincidental that this fluorescence layer was located between the Subtropical Subsurface Water¹⁷ salinity maximum, and the Intermediate Water salinity minimum.

This suggests that these fluorescent particles may have originated in the highly productive equatorial surface waters¹⁸, and were advected into the region. A low density of filter feeders in the oxygen-minimum zone is consistent with the bioluminescence profiles, because the preservation of this particle layer requires low grazing rates compared with the rate of lateral particle injection from the source region.

These hypotheses will be tested on a subsequent cruise by obtaining beam attenuation and fluorescence profiles over a larger geographical range. Nevertheless, the present data demonstrate some exciting potential uses of bio-optical sensors in the deep sea.

This research was performed as part of the VERTEX project supported by the NSF, grant no. OCE 79-26797. We thank Sheila Baldrige, Don Eliason, Chris Pelletier, Richard Reaves and Nancy Greene Teas for assistance, and George Knauer and Stevens Tucker for helpful discussions.

Received 3 November 1982; accepted 13 January 1983.

1. Herman, A. W. & Denman, K. L. *Deep-Sea Res.* **24**, 385-397 (1977).
2. Yentsch, C. S. & Menzel, D. W. *Deep-Sea Res.* **10**, 221-231 (1963).
3. Lorenzen, C. J. *Deep-Sea Res.* **13**, 223-227 (1966).
4. Strickland, J. D. H. *Deep-Sea Res.* **15**, 225-227 (1968).
5. Kiefer, D. A. *Mar. Biol.* **22**, 263-269 (1973).
6. Loftus, M. E. & Seliger, H. H. *Chesapeake Sci.* **16**, 79-92 (1975).
7. Duursma, E. K. in *Optical Aspects of Oceanography* (eds Jerlov, N. G. & Nielsen, E. S.) 237-256 (Academic, New York, 1974).
8. Jerlov, N. G. *Marine Optics* (Elsevier, Amsterdam, 1976).
9. Kiefer, D. A. & Austin, R. W. *Limnol. Oceanogr.* **19**, 55-64 (1974).
10. Tett, P. B. & Kelly, M. G. *Oceanogr. mar. Biol. a. Rev.* **11**, 89-173 (1973).
11. Strickland, J. D. H. & Parsons, T. R. *A Practical Handbook of Seawater Analysis* (Queen's, Ottawa, 1968).
12. Kelly, M. G. & Tett, P. B. in *Bioluminescence in Action* (ed. Herring, P. J.) 399-417 (Academic, London, 1978).
13. Hardy, A. C. & Kay, R. H. *J. mar. Biol. Ass. U.K.* **44**, 435-484 (1964).
14. Setser, P. J., Guinasso, N. L. Jr & Schink, D. R. *J. mar. Res.* **40**, 453-471 (1982).
15. Broenkow, W. W. & Krenz, R. T. *Moss Landing Marine Laboratories Tech. Publ.* 82-1 (1981).
16. Silver, M. W. & Alldredge, A. L. *J. mar. Res.* **39**, 501-530 (1981).
17. Wyrtki, K. *Int. J. Oceanogr. Limnol.* **1**, 117-147 (1967).
18. Love, C. M. (ed.) *EASTROPAC Atlas*, 8 (National Oceanic and Atmospheric Association, Washington DC, 1974).

The time to extinction of a colonizing propagule of lizards increases with island area

Thomas W. Schoener

Department of Zoology, University of California, Davis, California 95616, USA

Amy Schoener

Department of Oceanography, University of Washington, Seattle, Washington 98195, USA

The question of what causes species extinction and how to prevent it is a crucial one for both basic and applied ecology. Three kinds of island-biogeographical field studies have made particularly important contributions towards this problem. First, a variety of bird studies suggest for small populations a negative relation between extinction time and population size; smaller islands typically have smaller populations (see ref. 1 and H. L. Jones, J. M. Diamond and M. Gilpin, manuscript in preparation). Second, in a study on arthropods², experimentally reducing the area of already saturated mangrove islands caused numbers of species to decrease, presumably because of reduced population sizes. Third, in a study of mammals³ in which small, variably sized populations of colonists (propagules) were experimentally introduced onto islands, larger propagules survived longer. In the present study with lizards, both island area and propagule size were varied experimentally in an attempt to separate their effects. The results show: (1) for a given species, time to extinction increased with island area in a monotonic fashion; (2) above a certain island area, rapid colonization occurred; below that threshold, propagules went extinct quickly; and (3) propagule size had no effect.

Naturally occurring species of lizards on small Bahamian islands typically show a highly regular relation to island area: each species has a rather precise threshold, above which it always occurs and below which it never occurs^{4,5}. To investigate this pattern, and to elucidate the general problem of what determines minimal areal requirements of species, we placed propagules of lizards on 30 islands that had no lizards. These islands spanned 4 orders of magnitude in area, the largest island being just short of the 0-to-1 species threshold. By varying propagule size and island area, and by following the time course of extinction, we wished to observe or infer the major factors causing extinction of our experimental propagules. Our expectation, based on a previous attempt by Levins and Heatwole⁶ to introduce lizards and upon the above pattern, was that most if not all introductions would fail fairly rapidly.

Islands in three localities, spanning 1,000 km, were chosen as experimental sites. Two localities (Abaco, Exumas) have *Anolis sagrei* as the lizard naturally occurring on 1-species islands⁴⁻⁵; this species appears from distributional data to be a good colonizer⁷. One (Acklins) has *Leiocephalus punctatus* as the solitary species⁴⁻⁵. All introductions were simultaneous within a site, and all but one used the solitary species appropriate to that site (Table 1). Small (three adult females, two adult males) and large (twice as many of each) propagules were used; their sex ratio approximated that in high-density, large-island populations⁸. To investigate the effect of propagule size, we chose eight pairs of islands to be as alike in area and vegetation as possible (Table 1). We introduced a small propagule to a randomly chosen member of each matched pair, and a large propagule to the other. Lizards were collected from that main island nearest to the experimental islands; in no case did the distance exceed 16 km. Lizards were collected the day or night immediately preceding introduction and stored in a dark, cool place until transport to minimize physiological stress. We monitored most populations: (1) twice during the 10-day period immediately following introduction; (2) at two 6-month intervals after introduction; and (3) yearly thereafter.

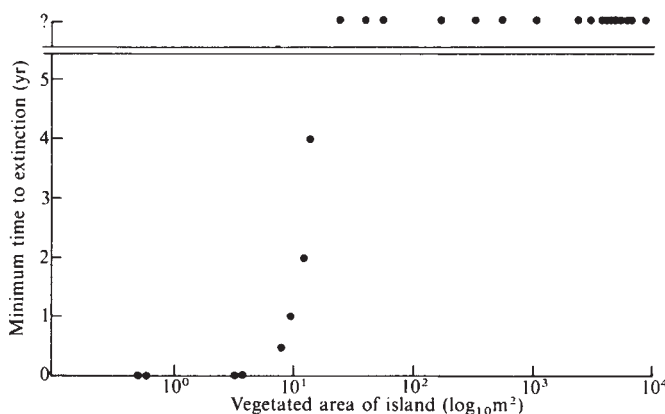
We defined extinction in two ways. 'Total extinction' has occurred when all individuals have disappeared from an island. 'Reproductive extinction' has occurred when total extinction has occurred, or when only males survived on an island, or when only females were found for two consecutive observation periods with no evidence of reproduction (females, during the time we observed them, may have eggs). Clearly, reproductive extinction is the more important quantity so far as population persistence is concerned.

The major result of our study is that, within a species, time to reproductive extinction increases monotonically with vegetated area of an island (Table 1). Figure 1 illustrates this result for *A. sagrei*. Time to total extinction, except for one island (N11 at 6 months; a natural recolonization by a single female during the third year appears also to have occurred here), also increases monotonically with vegetated area (though there are two more ties). Times to extinction are not monotonically related to total island area, but the relationship is still extremely strong, as the two measures of area are highly correlated (Table 1).

In our study, propagule size had no relation at all to time to extinction: large and small propagules fall along the same curve (Fig. 1). However, the larger species, *L. punctatus*, becomes extinct sooner for a given island area than does the smaller species, *A. sagrei*. The former species has a larger minimal home range and occurs at substantially lower natural densities than does *A. sagrei* (ref. 8 and C. H. Stinson and T.W.S., manuscript in preparation) so that this result is unsurprising⁹.

As only adults were introduced, we could ascertain from frequent observation without capturing lizards and thereby disturbing the potentially quite delicate stages of initial colonization whether reproduction had occurred on each island. On all but the seven smallest islands, some reproduction occurred. On the largest islands, reproductive rates were explosive. For example, White Bay Cay was seeded with 10 *A. sagrei* and the following year had an estimated 98 individuals. Moreover, survival rate on this island, 80% in the first year, is the highest ever recorded for an *Anolis* lizard¹⁰ and the second highest for any lizard species anywhere¹¹.

Using data from the largest island (White Bay Cay), we estimate the instantaneous death rate, μ , as 1.39 per yr and, reproductive rate, r as 2.12 per yr (methods according to Crowell³). The instantaneous birth rate, λ , is $r + \mu = 3.51$. Presumably μ and λ determined for the largest island come close



to the exponential (no effect of carrying capacity) rates in the demographic-stochastic models of MacArthur and Wilson¹² and Goel and Richter-Dyn¹³. In these models, for a variety of assumptions concerning the mathematical form of population limitation, and so long as populations are sufficiently below their carrying capacities, extinction probability, P_E , is approximately $(\mu/\lambda)^n$, where n is the number of individuals (here females) in the propagule. Substituting $n=3$ or $n=6$ gives $P_E=0.062$ or $P_E=0.004$, respectively. Thus in retrospect, because of the enormous reproductive potential of our lizards on apparently highly favourable islands, neither-sized propagule we used was at all likely to go extinct; even if $n=1$, colonization has a better-than-even chance.

The same models also predict the existence of a critical population size, n_c , above which extinction is very unlikely. Hence, if an island's carrying capacity, K , exceeds n_c , populations should reach K and persist there for a very long time, barring environmental change. The threshold n_c is largely independent of K and for our values of μ and λ works out to 7–10 females or less than 15–20 individuals in all. This estimate might be a little low, as μ and λ were calculated from the island presumed to be most favourable. However, actual population counts on the smallest islands still having lizards are remarkably close to those predicted (1978–80 populations were estimated by the regression technique of Tanaka¹⁴, 1981 Acklins populations by mean Lincoln index, other populations by the multi-way-contingency-table method of Fienberg¹⁵ as adapted for lizards by Heckel and Roughgarden¹⁶). In the Acklins site, South White Bay Cay had 20 (95% confidence interval 13–27) *A. sagrei* in 1982; counts for *L. punctatus* on North White Bay for 1978–82 are 12 (10–13), 12 (10–27), 14 (13–16), 17 (15–19) and 11 (11–11). In the Exuma site, Cay 3 had 27 (13–41) individuals in 1982, and Cay 15 had 18 (10–26) and 33 (17–49) individuals in 1981–82. Larger islands have counts ranging to

97 (93–103) individuals for *L. punctatus* and 490 (427–553) individuals for *A. sagrei*. So far as we can tell, populations are stable (for example, in addition to the above data, populations on White Bay were 307 (180–1052), 407 (366–1274) and 391 (357–486) individuals in 1979, 1980 and 1982, and 3 other islands showed an approximately 20% increase from 1981 to 1982). Hence, our populations now appear to be close to K . On islands where lizards have gone extinct, populations never exceeded initial (propagule) sizes, and extinction occurred in periods ranging from 5–11 days to 3 yr. Thus the existence of a 'safe' threshold population size, corresponding to a safe island area, is supported by our data.

Our results strongly support the above-cited models of island colonization, and they are also consistent with similar approaches in epidemiology^{1,13,17}. However, there are two anomalies. First, why did lizards on the four smallest islands all disappear within 5–11 days (in one case, this involved a propagule size of 10)? Second, experimentally established populations are stable on islands down to two orders of magnitude smaller than islands naturally having lizards, so why did those islands not have lizards in the first place? The surprising success of our introductions is in contrast to the failures in the only previous such study with lizards, that of Levins and Heatwole⁶.

Rapid extinctions on the smallest islands could not have involved starvation, as the subject species can live without food (or water) in captivity for at least 2 weeks. Predation is unlikely; few lizard-eating predators occur on any small island in the study sites, and the islands in question have none whatsoever. We cannot imagine transient predators, such as semi-aquatic birds, being able to wipe them out so quickly as hiding places were abundant. We believe either that the lizards became physiologically stressed, went into holes and never emerged, or left the islands voluntarily. In support of the latter, we have

Table 1 Introduction experiments onto empty islands

Island	Total area ($\times 10^{-2}$ m ²)	Vegetated area ($\times 10^{-2}$ m ²)	Introduced species	Propagule size	Under 14 days: status (days)	Time interval					
						5–6 months	1st yr	2nd yr	3rd yr	4th yr	5th yr
Acklins											
A1	0.72	0.58	L	5	0(8), 0(9)	0	0	0	0	0	0
A2	1.55	0.78	L	5	X(9)	X	0	0	0	0	0
N. White Bay	4.88	1.62	S	5	M(9)	B-R	B	B	B	B	B
S. White Bay	3.58	1.79	L	5	B(3), X(9)	X-R	B	B	B	B	B
W. Money	4.84	4.36	L	11	B(9)	X-R	B	B	B	B	B
Barracuda Rock	10.1	10.1	L	5	X(9)	X-R	B	B	B	B	B
Exumas											
E5	3.27	0.006	S	10	0(4), 0(10)	0	0	0	0	0	0
E17	4.41	0.03	S	5	0(5), 0(11)	0	0	0	0	0	0
E4	2.81	0.04	S	5	0(4), 0(10)	0	0	0	0	0	0
E22'	2.33	0.08	S	5	B(5), F(10)	F	0	0	0	0	0
E6	1.76	0.09	S	5	F(4), 0(10)	F-R	F	M	0	0	0
E20	4.20	0.12	S	5	F(5), F(11)	B-R	M	F	0	0	0
E8	2.46	0.13	S	5	B(5), B(9)	B-R	B	B	B	F	F
E3	3.38	3.05	S	5	B(4), F(10)	B-R	B	F	B	B	B
E12	6.35	5.08	S	10	B(5), B(11)	B-R	B	F(B?)	B	B	B
E24	11.1	10.0	S	10	B(4), M(10)	B-R	B	B	B	B	B
E16	24.2	21.7	S	10	M(5), F(11)	B-R	B	B	B	B	B
E21	30.7	27.6	S	5	B(4), F(10)	B-R	B	B	B	B	B
E18	40.3	40.3	S	5	M(5), F(11)	B-R	B	B	B	B	B
E19	40.3	40.3	S	10	B(5), F(11)	B-R	B	B	B	B	B
E15	40.3	40.3	S	5	M(5), M(11)	M-R	B	B	B	B	B
E13	44.3	44.3	S	5	F(5), F(11)	B-R	B	B	B	B	B
E11	48.4	48.4	S	5	M(5), F(11)	B-R	B	B	B	B	B
Flat Rock	56.5	56.5	S	10	B(5), F(11)	B-R	B	B	B	B	B
Andrew	56.5	56.5	S	5	B(10)	B-R	B	B	B	B	B
White Bay	80.6	80.6	S	10	M(10)	B-R	B	B	B	B	B
Abaco											
N11	0.26	0.005	S	5	M(9)	M	0	0	X	F	
N7	0.22	0.22	S	5	B(9)	B-R	B	B	B	B	
N13	0.35	0.35	S	5	F(9)	B-R	B	B	B	B	
N9	0.51	0.51	S	5	B(9)	B-R	B	B	B	B	

L, *Leiocephalus punctatus*; S, *Anolis sagrei*; R, juveniles produced; M, only males found; F, only females found; B, both sexes found; X, species occurs; 0, species absent.

shown elsewhere¹⁸ that for smaller islands lizards frequently jump into the sea, possibly to swim or drift to another island.

Successful colonization of most islands not having lizards naturally shows that chronic factors, such as desiccation of eggs or adults, or starvation, cannot be agents of extinction. Perhaps populations will eventually overeat their food and crash; however, severe oscillations are theoretically unlikely in systems such as ours where much food 'drifts' in from elsewhere¹⁹, and our numbers give no evidence of them. Several lines of evidence (refs 4, 5, 19, manuscript in preparation), of which the above results comprise only one, imply that minimal areas supporting lizard populations reflect 'bottlenecks' caused by periodic catastrophes, in this case especially destructive hurricanes. Our hypothesis, aspects of which we are now testing experimentally, is that hurricanes cause not only much short-term mortality, but also long-term devastation of the vegetation and concomitant resource shortage. During the aftermath of a hurricane, populations of lizards go extinct on islands that, once their vegetation has recovered, could support them, sometimes lavishly. For this explanation to hold, immigration during favourable times must be virtually nil; indeed, we also hypothesize that immigration is pulse-like in this system, occurring during and immediately following a hurricane, when lizards are especially likely to be in the water. In short, we suggest that present-day distributions of lizards on small islands are in several senses the high-water marks of past disasters.

We thank C. Bayer, C. Stinson and S. Rissing for help with the censuses, C. Stinson and M. Turelli for reading the manuscript, J. Roughgarden, D. Heckel and S. Syrjala for programs, and the NSF for support (DEB 78-22798, DEB 81-18970).

Received 4 October 1982; accepted 11 January 1983.

1. Williamson, M. *Island Populations* (Oxford University Press, 1981).
2. Simberloff, D. *Ecology* **57**, 629-648 (1976).
3. Crowell, K. L. *Am. Nat.* **107**, 535-558 (1973).
4. Schoener, T. W. & Schoener, A. J. *Amin. Ecol.* **52**(1) (1983).
5. Schoener, T. W. & Schoener, A. J. *Anim. Ecol.* **52**(1) (1983).
6. Levins, R. & Heatwole, H. *Ecology* **54**, 1056-1064 (1973).
7. Williams, E. E. Q. *Rev. Biol.* **44**, 345-389 (1969).
8. Schoener, T. W. & Schoener, A. J. *Anim. Ecol.* **49**, 19-53 (1980).
9. Schoener, T. W. *Proc. 16th int. orn. Congr.*, 629-642 (1976).
10. Schoener, T. W. & Schoener, A. *Oikos* **39**, 1-16 (1982).
11. Turner, F. B. in *Biology of the Reptilia* Vol. 7 (eds Gans, C. & Tinkle, D. W.) 157-264 (Academic, London, 1977).
12. MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton University Press, 1967).
13. Goel, N. S. & Richter-Dyn, N. *Stochastic Models in Biology*, Ch. 4 (Academic, New York, 1974).
14. Seber, G. A. F. *The Estimation of Animal Abundance*, 145 (Griffin, London, 1973).
15. Fienberg, S. E. *Biometrika* **59**, 591-603 (1972).
16. Heckel, D. G. & Roughgarden, J. *Ecology* **60**, 966-975 (1979).
17. Black, F. L. J. *theor. Biol.* **11**, 207-211 (1966).
18. Schoener, T. W. & Schoener, A. in *Festschrift for Ernest Williams* (eds Rhodin, A. & Miyata, K.) (Special Publ. Museum Comparative Zoology, Harvard University, 1982).
19. Schoener, T. W. *Theor. Pop. Biol.* **4**, 56-84 (1973).

Polyandry, cloaca-pecking and sperm competition in dunnocks

N. B. Davies

Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

A variety of behaviours adopted by males before and after copulation serve to increase paternity. The most spectacular examples occur in insects where males increase their own chances of fertilizing the female's eggs by mate guarding, mating plugs and even removal of other males' sperm^{1,2}. Here, sperm competition is described for a small European passerine bird, the dunnock (*Prunella modularis*), where females are often mated simultaneously to two males^{3,4} and where there is an elaborate pre-copulatory display^{5,6}. It is shown that, during this display, the male stimulates the female to eject sperm before he himself copulates. The display is most intense where there is a high probability that another male has recently mated with the female.



Fig. 1 *a*, The male first stands behind the female and pecks her cloaca. Mean duration of display was 50.2 s, range 5-120 s. Mean number of pecks was 27.9, range 0-118 pecks; 74 displays observed. *b*, Then he copulates. Copulation itself is very brief, the male appears to jump over the female, cloacal contact lasting for just a fraction of a second. (Drawing by John Busby).

The study site is the Cambridge University Botanic Garden where, at the start of the breeding season in 1982, there was a colour-ringed population of 48 males and 31 females. The male-biased sex ratio is probably a consequence of greater female mortality during winter. Females are smaller than males, subordinate at feeding sites and more likely to die or leave temporarily when competition for food is intense. The mating combinations on some of the breeding territories were very complicated but the commonest situations were either a monogamous pair (10 territories) or a trio with two males (not close relatives) associating with the same female (15 territories). In trios, one of the males, the alpha male, was dominant to the other, the beta male, both at feeding sites and over access to the female. Alpha males (mean wing length 70.5 mm) were significantly larger than beta males (mean 69.3 mm, $P < 0.02$).

Monogamous males and alpha males of trios began to follow the female closely, mainly staying within 5 m of her, 4 to 5 days before she laid her first egg, similar to the onset of mate guarding in other species⁷. Neighbouring males often encroached onto the territory and attempted to copulate with the female but by far the most intense competition occurred between the males of a trio. Here the beta male constantly approached the female and the alpha male spent up to 30 min per h chasing him away. There were often fights and in two cases the beta male was blinded in one eye and subsequently died.

Mate guarding behaviour was studied in detail for 25 breeding attempts, including first and second broods in 1981 and 1982 (data from 19 different territories). The female on each of these territories was watched continuously for 1-3 h daily during the mating period. In three of the four breeding attempts by monogamous pairs, only the resident male was seen to mate while in the other case a trespassing neighbour also mated with the female. Of the 21 breeding attempts by trios, only the alpha male was seen to copulate in 12 cases, while in the other 9 both alpha and beta male copulated. Beta males sometimes managed to approach the female and copulate without the alpha male noticing, for example when he was busy feeding or out of sight on the other side of a bush. On other occasions the alpha male lost the female temporarily after chasing the beta male. If the beta male was then the first to find the female again he was able to mate with her unhindered.

In most passerine birds the preliminary display immediately before copulation is very brief. In dunnocks by contrast, it is an extraordinarily elaborate affair. The female crouches, fluffs up her body feathers, shivers her wings and quivers her tail which is raised to expose the cloaca. The male hops from side to side behind her and pecks her cloaca for up to 2 min before he copulates (Fig. 1). Monogamous males and both alpha and beta males of trios pecked the female's cloaca before they mated.

During the pecking by the male, the female's cloaca becomes pink and distended and from time to time makes strong pumping movements. As the cloaca contracts the female jerks her abdomen downwards and sometimes is seen to eject a small pale mass. (There is one previous report⁸ of a female ejecting a 'small drop of fluid' during the cloaca contractions.) A small droplet was found on three occasions by careful search of the ground where a male and female had been displaying, and examination of the droplets through a microscope revealed them to be sperm masses (Fig. 2). One of the sperm masses was extruded together with a small amount of faecal material. On 11 other occasions the female extruded larger pellets during the pre-copulatory display and these were all confirmed to be faecal pellets only. It was impossible to see whether the female ejected something during every display and because the sperm droplets were so small it is possible that others were deposited but not found. The male sometimes pecked at the ejected mass and in all five cases where the display sequence was seen in close detail, he copulated with the female immediately after she had produced the deposit.

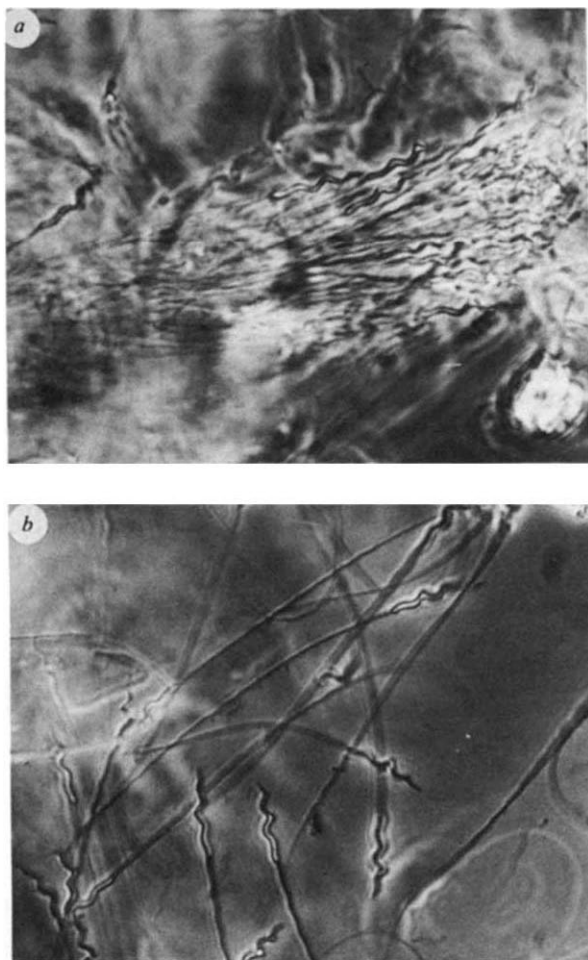


Fig. 2 *a*, Part of a sperm mass ejected by the female during cloaca pecking by the male. The semen of passerine birds is a small glutinous droplet, with little accessory fluid, containing bundles of sperm, which have helical heads¹³. *b*, Some individual sperm on the edge of the mass.

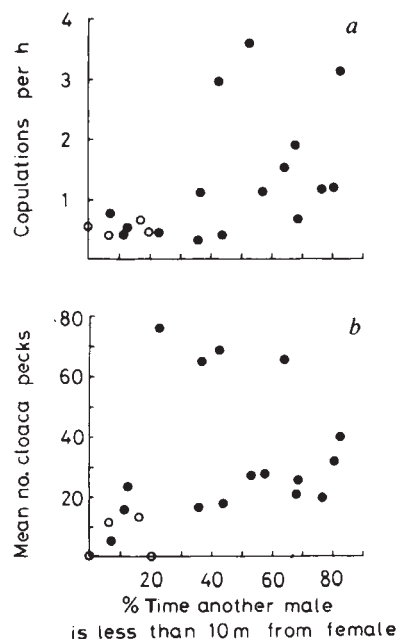


Fig. 3 *a*, Males copulate more frequently when they are with the female (less than 10 m) the greater the proportion of time another male spends close to her. Open circles are data from 4 monogamously paired males where 'other males' are trespassing neighbours. Closed circles are 16 alpha males of trios where most of the interference is by beta males resident on the territory (Spearman rank correlation on all points is 0.627, $P < 0.01$; on alpha males data only, it is 0.582, $P < 0.05$). *b*, Males also peck the female's cloaca more each time before they copulate the more time another male spends close to her (Spearman rank correlation on all points is 0.598, $P < 0.01$; on alpha male data only, it is 0.300, NS).

In birds, it is thought that the male ejaculates sperm into the female's vagina which is a short muscular duct joining the cloaca⁹. There are sperm-host glands in the utero-vaginal junction where the sperm is stored for some time before it is transported up to the top of the oviduct where fertilization takes place immediately after ovulation, about 24 h before the egg is laid⁹⁻¹¹. The egg has to be fertilized during its brief sojourn in the infundibulum, a period of only 15–30 min. It is unlikely that the female can be certain of copulating during this brief period which presumably explains why sperm is stored and then moved up the oviduct just before ovulation so as to be available at the critical time for fertilization¹¹.

One hypothesis for the cloaca-pecking by the male is that it stimulates the female to eject sperm likely to have been inseminated by another male. Any strong contractions in the female's urogenital tract would presumably eject faeces as well if any were in the cloaca at the time. In accord with this sperm competition hypothesis, a male both copulates more frequently and does more pecking in each pre-copulatory display the more time another male spends near his female (Fig. 3). In trios the proportion of time a beta male spends within 10 m of the female is negatively correlated with the alpha male's wing length (Spearman rank correlation -0.776 , $P < 0.01$, $n = 13$), which suggests that larger alpha males are better able to monopolize exclusive access to the female and keep beta males at bay. Monogamous males are even larger than alpha males of trios and are able to evict beta males out of their territories altogether (manuscript in preparation).

One of the presumed costs of a long pre-copulatory display is a high probability that a male will be interrupted by another male before he actually copulates. Fifty out of 125 copulation attempts by alpha males were interrupted during the cloaca-pecking by a beta male's arrival, whereupon the alpha male left the female to chase him. Fourteen out of 38 attempts by beta males were interrupted by the alpha male. The advantages of cloaca-pecking must be considerable to offset this cost. The

details of sperm competition are not known for any wild bird but one possibility is that if the female's sperm store becomes full then the only chance a male has of putting sperm in himself is to remove sperm first to make room. Even if the store is not full, however, it may pay a male to risk the cost of interruption in order to remove sperm which might not be his and insert sperm which he can be certain is his own.

It would be difficult to tell from field observations exactly how often a female ejects sperm but it is interesting to speculate on the possible advantage to her of doing so. One possibility is that it is a mechanism by which she increases the chances that her clutch of eggs is fertilized by more than one male. Females actively attempted to escape the guarding of the alpha male and often approached the beta male to solicit copulations. In all seven cases where both alpha and beta male copulated with the female and chicks were hatched successfully, both males fed the brood. By contrast, the beta male never fed the brood in the five cases where chicks hatched and only the alpha male had mated. Nestlings fed by two males get more food, fledge at a heavier weight and survive better than those fed by just one male (manuscript in preparation). Females can therefore increase their reproductive success if they get both males to fertilize their clutch or indeed persuade them both that they have some paternity¹².

I thank S. M. Walters and P. Orriss for the facilities of the Cambridge University Botanic Garden, the SERC for financial support and T. R. Birkhead, B. Gupta, D. Harper, A. I. Houston, A. Lundberg and W. Westley for their help and comments.

Received 14 December 1982; accepted 2 February 1983.

1. Parker, G. A. *Biol. Rev.* **45**, 525-568 (1970).
2. Waage, J. K. *Science* **203**, 916-918 (1979).
3. Birkhead, M. E. *Ibis* **123**, 75-84 (1981).
4. Snow, B. & Snow, D. *Misc. Publ. Yamashina Inst.* (in the press).
5. Delamain, J. *Br. Birds* **23**, 19-20 (1929).
6. Harrison, C. J. O. & Binfield, F. G. *Bird Study* **14**, 192-193 (1967).
7. Birkhead, T. R. *Anim. Behav.* **30**, 277-283 (1982).
8. Sanderson, R. F. *Bird Study* **15**, 213 (1968).
9. Lake, P. E. in *Form and Function in Birds* (eds King, A. S. & McLelland, M.) 1-62 (Academic, London, 1981).
10. Gilbert, A. B. *Adv. Reprod. Physiol.* **2**, 111-179 (1967).
11. Howarth, B. in *The Oviduct and its Functions* (eds Johnson, A. D. & Foley, C. W.) 237-270 (Academic, London, 1974).
12. Stacey, P. B. *Am. Nat.* **120**, 51-64 (1982).
13. Humphreys, P. N. J. *Reprod. Fert.* **29**, 327-336 (1972).

Magnetic particles in the liver: a probe for intracellular movement

Peter Gehr*, Joseph D. Brain†, Steven B. Bloom & Peter A. Valberg

Department of Physiology, Harvard University School of Public Health, 665 Huntington Avenue, Boston, Massachusetts 02115, USA

Previous studies have used magnetic particles to estimate the viscosity of cell cytoplasm *in vitro*¹⁻⁴. Here we describe how magnetic Fe₃O₄ particles can be used to estimate non-invasively the motion of organelles in hepatic macrophages in intact animals. We report that when these particles are injected intravenously (i.v.), most are phagocytosed by hepatic macrophages (Fig. 1)⁵. When an external magnetic field is applied to the rabbit, these particles become magnetized and aligned. After removal of the field, the particles collectively produce a remanent magnetic field which can be measured at the body surface. This field decreases with time due to particle rotation (relaxation)^{6,7}. As the particles are contained in phagosomes or secondary lysosomes, we conclude that motions of these organelles are responsible for the particle rotation and relaxation.

* Present address: Anatomisches Institut, Universität Bern, Postfach 139, CH-3000 Bern 9, Switzerland.

† To whom correspondence should be addressed.

Submicrometric γ -Fe₂O₃ particles⁸ (2 mg suspended in 1 ml physiological saline) were injected i.v. into four adult male New Zealand White rabbits. For each measurement of relaxation, the rabbits were anaesthetized with 50 mg kg⁻¹ ketamine-HCl given intramuscularly and then placed in a 60-mT magnetic field for 5 min. This time is adequate to magnetize all the particles and align them with the external field. To measure the gradual decay of the remanent magnetic field, the rabbit was placed inside a cylindrical moly-permalloy shield and its xiphoid process (a reliable reference point for the liver) repeatedly presented to a fluxgate magnetometer probe (Magnetoscopy F 1.067; Foerster Instruments, Pennsylvania). This procedure was continued for 30 min. Magnetizations and relaxation measurements were performed at 2 h, and 1, 2, 3 and 4 days after particle injection. To prevent magnetic contamination of the animals, they were kept in non-ferrous cages and fed only fresh vegetables and water; commercial rabbit food was found to contain magnetic contaminants.

After the final measurement on day 4, each animal was killed and the amount of magnetic material in the liver, spleen, lung, heart and stomach was determined by magnetizing a tissue sample from each organ and measuring its remanent field with the fluxgate magnetometer. More than 90% of the retained γ -Fe₂O₃ was found in the liver, some in the spleen, but only minute amounts in other organs. Samples of liver and spleen

Table 1 Relaxation rate for the first minute of measurement, λ_0

Day	$\lambda_0(\text{min}^{-1})$
0*	$0.408 \pm 0.120^\dagger$
1	0.357 ± 0.067
2	0.341 ± 0.138
3	0.262 ± 0.044
4	0.264 ± 0.068

* Actual time of measurement was 2 h after injection.

† Decrease of λ_0 with time is significant at $P < 0.02$, by analysis of co-variance. The means are shown with their standard deviations.

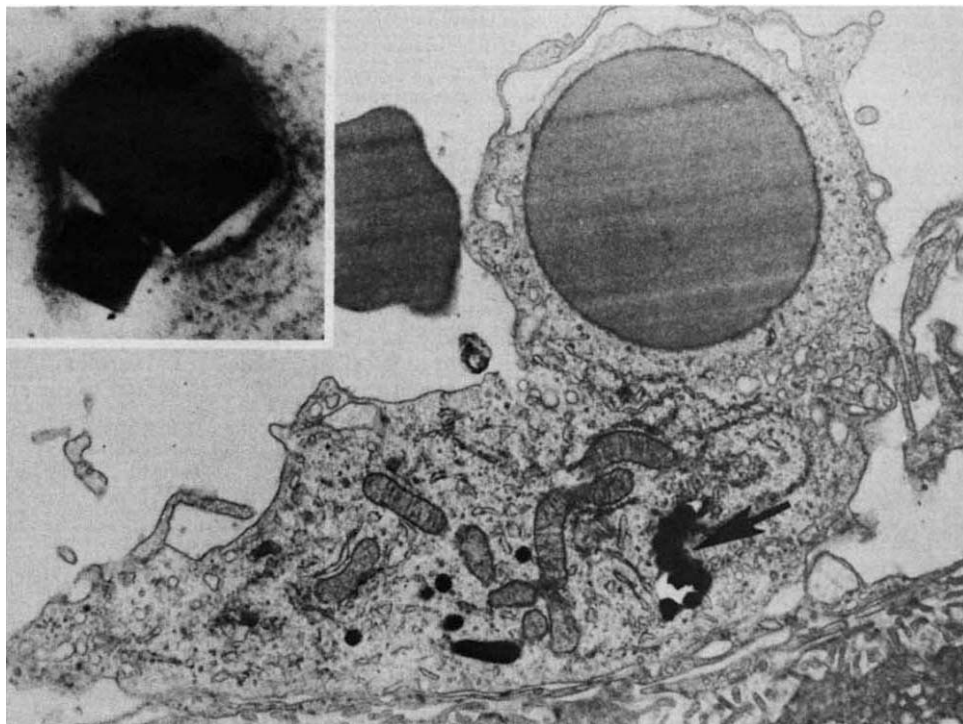
were examined by light and electron microscopy. Abundant iron oxide particles were observed in phagosomes and secondary lysosomes of Kupffer cells (Fig. 1); no particles were seen in any other cell type in the liver. Some particles were also found in splenic macrophages.

Figure 2 shows a relaxation measurement demonstrating the progressive decrease of the remanent field after removal of the external field. This reflects the gradual misalignment of the magnetized particles. As it took 10 s to move the animal from the magnetizing field to the shielded magnetometer probe, we estimated the initial remanent field strength, B_0 , by fitting the data to an exponential equation as suggested by Cohen *et al.*⁹ and extrapolating back to time zero. This equation, $B = B_0 e^{-\lambda_0 t}$ (where B is the field strength at time t) also allowed us to describe the rate of decay during the first minute, λ_0 .

During the 4 days of observation, B_0 (proportional to the amount of iron oxide present) did not change significantly; we conclude that there was negligible loss of particles from the liver during this time. The parameter λ_0 , however, decreased significantly (Table 1), showing that the process of particle misalignment gradually slowed. There was also a significant slowing during the remaining portion of each relaxation curve, but to a lesser degree. We also found that relaxation ceased shortly after the animal was killed on day 4 (Fig. 3); for about 6 min after the heart stopped beating, relaxation continued, then further particle rotation within the Kupffer cells stopped as indicated by a constant remanent field. When the animals were placed in a magnetizing field between 10 and 60 min after death, no relaxation was observed, although the same initial field (B_0) was produced.

This is the first demonstration of relaxation in an organ other than the lung. Relaxation of magnetic particles was first noted in the lungs of welders by Cohen⁶, and there have been additional reports of this phenomenon in other human^{7,9,10} and

Fig. 1 Hepatic macrophage from a rabbit injected with 2 mg of $\gamma\text{-Fe}_2\text{O}_3$. The cell, which has ingested magnetic particles (see arrow), is attached to the sinusoidal endothelium lying underneath. No other cell type in the liver showed associated particles. Note also the ingested red cell. $\times 9,100$. The inset shows that the crystalline $\gamma\text{-Fe}_2\text{O}_3$ particles are membrane bound. $\times 108,000$.



animal^{7,11,12} lungs. Cohen suggested that relaxation may be restricted to the lungs and that respiratory motion might be responsible⁷. Our results demonstrate that gross organ movements such as those produced by breathing are not essential for relaxation. Also, as relaxation persisted for several minutes after the heart stopped (Fig. 3), cardiac and arterial movements can be excluded. Other factors such as cell locomotion, cytoplasmic motion or thermal agitation remain as possible mechanisms. Of these, we believe that cytoplasmic motion is the most likely force responsible for particle rotation. Cell locomotion seems unlikely to contribute because Kupffer cells are an integral part of the sinusoidal endothelium (Fig. 1). As the injected particles were confined within phagosomes and secondary lysosomes in the Kupffer cells (Fig. 1), it seems probable that motions of particles within these organelles or the motions of the organelles themselves were involved. Do particles within phagosomes rotate by brownian movement? We believe not, as thermal agitation should continue after death. At 6 min post mortem, relaxation ceased, although body temperature was unchanged (Fig. 3). When magnetized after death, no relaxation was observed.

This was a consistent finding in all experimental animals. Thus, rotation of phagosomes and secondary lysosomes, both of which contain particles, is the most probable driving force for relaxation. Such organelle motion, due to cytoplasmic movements, is probably diminished by lack of oxygen and nutrients.

There is increasing evidence that organelles are not objects floating freely in the cytoplasm. Connections between organelles and microfilaments, a part of the cytoskeletal system, have been described^{13,14}; such attachments to the cytoskeleton should not only restrict random brownian movement of organelles but should also facilitate directed movement. Recently, Wolosewick and Porter¹⁵ have described a microtrabecular lattice, an even finer meshwork of very fine, short filaments which pervades the entire cytoplasm and connects organelles with larger structural elements. However, the existence of this network is still controversial; if real, it might also influence organelle movements¹⁶.

Processes requiring organelle motion such as cell secretion and formation of secondary lysosomes are not haphazard; rather, these events seem to be highly orchestrated processes

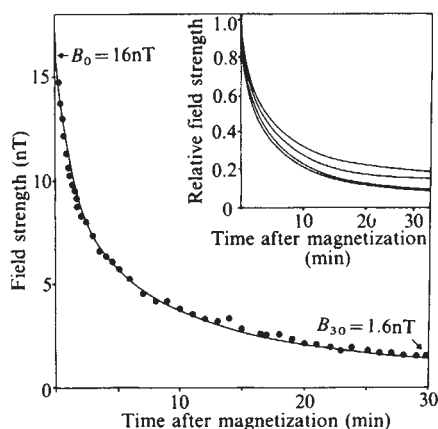


Fig. 2 Relaxation curves measured at 2 h post-injection. Individual readings (a single presentation to the fluxgate magnetometer) are indicated by the closed circles and are shown for one animal. B_0 is the strength of the magnetic field in nT when extrapolated to zero time. The inset shows the curve of the same rabbit and those of three additional ones also examined at 2 h post-injection, demonstrating the consistency of measurements.

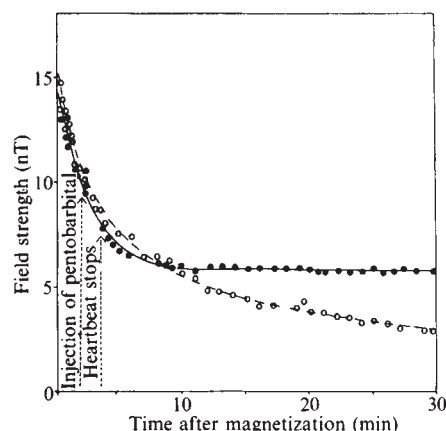


Fig. 3 Relaxation curve of a rabbit measured 4 days post-injection. During the first measurement, the animal was alive (open circles); 2 min after the onset of relaxation during the second measurement (closed circles), the rabbit was given a lethal injection of pentobarbital which stopped the heart beat in 2 min. Note that relaxation ceased ~ 6 min later.

involving the cytoskeletal and microtrabecular elements. We believe that contractions mediated by these structures cause the phagosomes and secondary lysosomes to move^{14,16,17}. These organelles would be rotated by contracting filaments much as puppets are moved by their strings.

The magnitude of organelle motion might be influenced by three factors: (1) the contractile forces available to rotate organelles; (2) the viscosity of the surrounding cytoplasm which would impede organelle rotations; and (3) the size of the organelles. The observed decrease in relaxation rate with increasing time after injection of the magnetic particles (Table 1) could reflect changes in these factors. Slowing of relaxation rate may reflect progression of particles from one part of the cell to another; different regions may have different rheological properties due to the amount and type of microfilaments present

and the extent of cross-linking^{18,19}. Increasing immobility of secondary lysosomes due to their increase in size by fusion with each other also seems possible.

We conclude that intracellular mechanisms are responsible for particle rotation which in turn produces relaxation. Contractions of the cytoskeleton causing phagosomal or secondary lysosomal motion may enhance relaxation, whereas formation of cross-links may restrain it. External measurements of retained magnetic particles could serve as a new and useful probe of such processes in phagocytic cells *in situ*.

We thank Dr Iku Nemoto for help in conducting the experiments, Dr Henry Feldman for statistical analysis, and Dr David Cohen for his advice regarding measurements of magnetic fields. We also thank Ms Antonia P. Velazco and Mr Leslie Phillips for typing the manuscript.

Received 28 July; accepted 21 December 1982.

- Heilbronn, A. L. *Jber. wiss. Bot.* **61**, 284–338 (1922).
- Crick, F. H. C. & Hughes, A. F. W. *Expl Cell Res.* **1**, 37–80 (1950).
- Yagi, K. *Expl Biochem. Physiol.* **3**, 73–91 (1961).
- Hiramoto, Y. *Symp. Soc. exp. Biol.* **22**, 311–327 (1968).
- Hampton, J. C. *Acta anat.* **32**, 262–291 (1958).
- Cohen, D. *Science* **180**, 745–748 (1973).
- Cohen, D. *Natn. tech. Inf. Serv. MIT/FBNML-78/1* (1978).
- Valberg, P. A. & Brain, J. D. *Am. Rev. resp. Dis.* **120**, 1013–1024 (1979).
- Cohen, D., Nemoto, I., Kaufman, L. & Arai, S. F. *IEEE Biomed. Engng* (submitted).

- Cohen, D., Arai, S. & Brain, J. D. *Science* **204**, 514–517 (1979).
- Halpern, M., Williamson, S. J., Spektor, D. M., Schlesinger, R. B. & Lippmann, M. *Expl Lung Res.* **2**, 27–35 (1981).
- Brain, J. D. & Valberg, P. A. *Am. Rev. resp. Dis.* **120**, 1325–1373 (1979).
- Agua, A. P. J. *ultrastruct. Res.* **74**, 175–182 (1981).
- Moore, P. L., Bank, H. L., Brissie, N. T. & Spicer, S. S. *J. Cell Biol.* **71**, 659–666 (1976).
- Wolosewick, J. J. & Porter, K. R. *J. Cell Biol.* **82**, 114–139 (1979).
- Luby, K. J. & Porter, K. R. *Cell* **21**, 13–23 (1980).
- Rebhun, L. J. in *Primitive Motile Systems in Cell Biology* (eds Allen, R. D. & Kamiya, N.) 503–525 (Academic, New York, 1964).
- Taylor, D. L. & Condeelis, J. S. *Int. Rev. Cytol.* **56**, 57–144 (1979).
- Temminck, J. H. M. & Spieles, J. J. *Cell Sci.* **41**, 19–32 (1980).

T-cell-mediated enhancement of host-versus-graft reactivity in mice fed a diet enriched in vitamin A acetate

Miroslav Malkovský, Andrew J. Edwards, Ruth Hunt, Lesley Palmer & Peter B. Medawar

Transplantation Biology Section, Clinical Research Centre, Watford Road, Harrow, Middlesex HA1 3UJ, UK

Retinol (vitamin A) and some of its derivatives have an important role in: (1) regulating growth, proliferation and differentiation of various tissues and (2) maintaining reproduction and visual function in man and higher animals^{1–5}. Vitamin A and retinoids are also known as potent immunoregulatory^{6–14} and antineoplastic agents^{15,16}. Their ability to increase reactivity to histoincompatible tissues is well documented^{7,8,17} but the mechanism of this action is unclear. Here we report that mice fed on an otherwise conventional diet supplemented with vitamin A acetate (VAA) respond to 10^5 semiallogeneic cells (a suboptimal dose) in a host-versus-graft (HvG) reaction, whereas mice on a conventional diet do not. It is possible to transfer this enhanced immune reactivity by injecting lymphoid cells from VAA-fed animals into those syngeneic mice maintained on the conventional diet. Using a positive selection technique, we demonstrate that the phenotype of the cell probably responsible for this phenomenon is $\text{Lyt } 1^+ 2^-$.

The sex- and age (6–9 weeks old)-matched mice used in these experiments were maintained on a conventional pelleted diet (Spratts Laboratory Diet 1, Spillers) *ad libitum*, with or without supplementary VAA (0.5 g per kg) in the form of stable gelatinized beadlets (Roche) as described in detail elsewhere^{17,18}. The rationale of this study was based on previous observations showing accelerated rejection of histoincompatible skin grafts and enlargement of the thymus and peripheral lymph nodes in mice treated with retinoids^{7,8,17}. Initially, we confirmed these results in our VAA diet system, demonstrating significantly enhanced reactivity to H-Y-incompatible skin grafts in VAA-fed animals; (CBA $\times$ C57BL/10ScSn) F_1 hybrid (CBF₁) females on the VAA diet rejected CBA male skin grafts substantially faster than control mice (R.H., L. P. and M. M., unpublished data). We then studied the effect of the VAA diet on the HvG

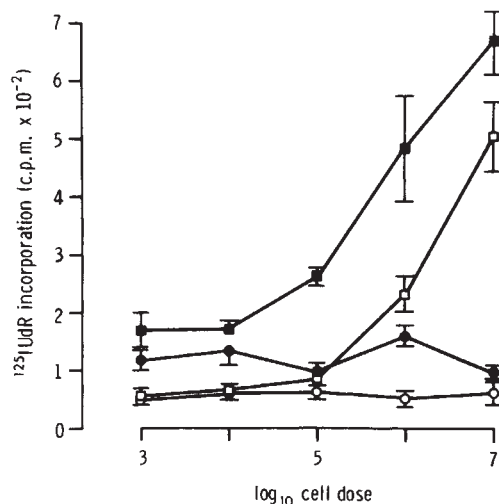


Fig. 1 Increased reactivity to alloantigens in VAA-fed mice. The abscissa indicates the $\log_{10}$ dose of viable syngeneic (circles) or semiallogeneic (squares) spleen cells injected into footpads of control (open symbols) and VAA-fed (close symbols) mice. The ordinate indicates ^{125}I UDr incorporation (c.p.m.) into the popliteal lymph nodes. The bars represent the standard errors of four determinations. Syngeneic and allogeneic stimulations were statistically compared by Student's *t*-test, with significance at $P < 0.05$. Doses of 10^6 and 10^7 semiallogeneic cells did, whereas doses of 10^3 and 10^4 semiallogeneic cells did not significantly increase DNA synthesis in the popliteal lymph nodes of both control and VAA-fed mice. However, a dose of 10^5 semiallogeneic cells significantly increased DNA synthesis in the popliteal lymph nodes of VAA-fed but not control mice.

Methods: Cell suspensions from male CBA and CBF₁ spleens were prepared using a fine mesh screen and erythrocytes were lysed by brief exposure to Tris-buffered ammonium chloride solution as described previously^{19,23}. Counts of viable cells were made in 0.05% Trypan blue. Cells suspended in a balanced salt solution (BSS) were injected subcutaneously into hind footpads. Syngeneic cells were injected into the right hind footpad and CBF₁ cells were injected into the left hind footpad of male CBA mice. Cell proliferation in popliteal lymph nodes was measured by following incorporation of ^{125}I UDr²⁴; 70 h after the footpad injections, mice were injected intraperitoneally with $10 \mu\text{Ci}$ of ^{125}I UDr (specific activity $\sim 5 \text{ Ci mg}^{-1}$; Amersham). Fifteen hours later, the ^{125}I UDr activity in popliteal lymph nodes was determined by using NE 1600 gamma counter (Nuclear Enterprises).

Table 1 Transfer of enhanced reactivity to alloantigens by lymphoid cells from VAA-fed mice

Experiment	Group	Responder mice (4 mice per group)	Numbers of intravenously injected syngeneic lymph node cells from:		¹²⁵ IUDr activity in popliteal lymph nodes stimulated with:	
			Conventional diet	VAA-enriched diet	Syngeneic cells (mean c.p.m. ± s.e.)	Semiallogeneic cells
<i>a</i>	A	Control CBA	—	—	21.8 ± 3.9	22.7 ± 3.2
	B	VAA-fed CBA	—	—	37.5 ± 5.6	155.3 ± 14.8
	C	Control CBA	5 × 10 ⁶	—	20.1 ± 3.7	20.6 ± 4.2
	D	Control CBA	10 ⁷	—	38.1 ± 7.2	39.6 ± 7.0
	E	Control CBA	—	5 × 10 ⁶	24.1 ± 5.2	128.0 ± 17.2
	F	Control CBA	—	10 ⁷	28.1 ± 5.5	141.2 ± 12.3
<i>b</i>	A	Control C57BL/10ScSn	—	—	17.5 ± 1.9	20.4 ± 2.5
	B	VAA-fed C57BL/10ScSn	—	—	20.9 ± 2.1	139.9 ± 16.3
	C	Control C57BL/10ScSn	10 ⁶	—	42.5 ± 18.4	50.2 ± 17.5
	D	Control C57BL/10ScSn	3 × 10 ⁶	—	30.7 ± 6.2	31.6 ± 2.3
	E	Control C57BL/10ScSn	9 × 10 ⁶	—	18.9 ± 4.4	23.9 ± 2.9
	F	Control C57BL/10ScSn	—	10 ⁶	32.5 ± 5.7	44.8 ± 4.4
	G	Control C57BL/10ScSn	—	3 × 10 ⁶	13.8 ± 4.9	141.5 ± 18.0
	H	Control C57BL/10ScSn	—	9 × 10 ⁶	28.2 ± 5.8	166.2 ± 13.6

Suboptimal doses of viable male CBF₁ and syngeneic spleen cells (10⁵ cells per footpad) were injected subcutaneously into hind footpads of CBA (expt *a*) or C57BL/10ScSn (expt *b*) male mice on the conventional or VAA-enriched diet. Immediately after, some of the mice on the conventional diet were injected intravenously with various doses of pooled syngeneic axillar, brachial and inguinal lymph node cells from control or VAA-fed male mice. Cell proliferation in the popliteal lymph nodes was assessed by determining ¹²⁵IUDr(5-[¹²⁵I]iodo-2'-deoxyuridine) incorporation as described in Fig. 1 legend. Syngeneic and allogeneic stimulations were statistically compared by Student's *t*-test. The *P* values for the groups B, E and F in *a* and groups B, G and H in *b* were <0.01; *P* for all other groups was >0.05.

Table 2 T cells expressing Lyt 1 but not Lyt 2 transfer the enhanced reactivity

Group	CBA responder mice (4 mice per group)	Intravenously injected cells from VAA-fed mice	Pretreatment of the intravenously injected syngeneic lymph node cells	Adherence to F(ab') ₂ anti-mouse plates	¹²⁵ IUDr activity in popliteal lymph nodes stimulated with:	
					Syngeneic cells (mean c.p.m. ± s.e.)	Semiallogeneic cells
A	Control	—	—	—	34.9 ± 9.8	40.3 ± 3.7
B	VAA-fed	—	—	—	21.2 ± 4.6	151.2 ± 12.6
C	Control	3 × 10 ⁶	Monoclonal anti-Thy 1.2 antibody	Adherent	13.2 ± 1.9	139.6 ± 16.5
D	Control	3 × 10 ⁶	Monoclonal anti-Thy 1.2 antibody	Nonadherent	29.4 ± 2.0	32.7 ± 7.0
E	Control	3 × 10 ⁶	Monoclonal anti-Lyt 1.1 antibody	Adherent	41.0 ± 8.3	160.8 ± 27.0
F	Control	3 × 10 ⁶	Monoclonal anti-Lyt 1.1 antibody	Nonadherent	51.4 ± 8.9	54.1 ± 6.6
G	Control	3 × 10 ⁶	Monoclonal anti-Lyt 2.1 antibody	Adherent	17.8 ± 1.3	20.5 ± 2.6
H	Control	3 × 10 ⁶	Monoclonal anti-Lyt 2.1 antibody	Nonadherent	17.6 ± 5.4	152.9 ± 11.0
I	Control	3 × 10 ⁶	—	Nonadherent	14.3 ± 2.0	141.5 ± 19.3

Nylon wool-purified²⁵ lymph node cells from VAA-fed female CBA mice were separated using monoclonal antibodies and a 'panning' technique of Mage *et al.*^{26,27} into marker-positive and marker-negative subpopulations^{19,28}. The F(ab')₂ fragment was prepared by pepsin digestion²⁹ of IgG obtained from rabbits immunized with mouse serum immunoglobulin²⁸, and was supplied by Dr G. L. Asherson. In brief, 9-cm Petri dishes (Sterilin) were coated with the F(ab')₂ antibody fragments²⁸. The cells were incubated with monoclonal antibodies (4 °C, 45 min) then washed and panned on the F(ab')₂ anti-mouse plates for 1 h at room temperature. The adherent cells were considered as marker-positive cells and the nonadherent cells as marker-negative cells^{19,26-28}. (None of the nylon wool column effluent cells untreated with monoclonal antibodies adhered to the F(ab')₂ anti-mouse plates.) Suboptimal doses of viable female CBF₁ and syngeneic spleen cells (10⁵ cells per footpad) were injected subcutaneously into hind footpads of female CBA mice on the conventional or VAA-enriched diet. Immediately after, some of the mice on the conventional diet were injected intravenously with 3 × 10⁶ marker-positive or marker-negative cells. Cell proliferation in the popliteal lymph nodes was assessed using ¹²⁵IUDr incorporation as described in Fig. 1 legend. Syngeneic and allogeneic stimulations were statistically compared by Student's *t*-test. The *P* values for the groups B, C, E, H and I were <0.01; *P* for all other groups was >0.05.

response of mice to alloantigens presented by viable CBF₁ spleen cells (Fig. 1). We found that whereas the dose of 10³ or 10⁴ CBF₁ cells was nonimmunogenic and the dose of 10⁶ or 10⁷ CBF₁ cells was immunogenic for all mice, the dose of 10⁵ viable semiallogeneic spleen cells was immunogenic for VAA-fed mice but not for mice on the conventional diet. A second set of experiments examined the possibility of transferring passively ('adoptively') this increased reactivity to alloantigens (Table 1). Syngeneic lymph node cells from VAA-fed mice injected intravenously at the time of administration of syngeneic and CBF₁ spleen cells into hind footpads led to an increased immune responsiveness in both CBA and C57BL/10ScSn mice. The fact that the effective dose of syngeneic lymph node cells transferred from VAA-fed mice was relatively low (~3 × 10⁶ cells per mouse) suggests that the 'acting' cell had a helper or inducer rather than effector function. Finally, we investigated the phenotype of this cell. We separated the lymph node cell

population from CBA mice on the VAA diet by using monoclonal antibodies and a 'panning' technique¹⁹ into marker-positive and marker-negative subpopulations for transfer (Table 2). The data obtained clearly indicate that the cell responsible for the increased immune responsiveness is a T cell expressing Lyt 1.1 but not Lyt 2.1 cell surface antigens. Furthermore, the expression of these cell surface antigenic determinants was quantified by a fluorescence-activated cell sorter (FACS) analysis, which revealed that VAA-fed mice had a significantly higher proportion of Lyt 1.1⁺ cells in their lymph nodes than control mice, whereas the percentage of Lyt 2.1 positive cells was unchanged (Table 3). Although this increase in relative numbers of cells having the Lyt 1 phenotype does not itself provide a satisfactory explanation for the phenomena observed, the results suggest that the functional changes might be accompanied by an increased expression of Lyt 1.

Table 3 Expression of Lyt 1.1 and Lyt 2.1 antigens by FACS analysis

	% Lymph node cells expressing Lyt 1.1	Lyt 2.1
Control mice	83.6 ± 1.1	24.4 ± 1.4
VAA-fed mice	90.1 ± 0.6	24.2 ± 0.7
Student's <i>t</i> -test	<i>P</i> < 0.001	Non significant difference

Immunofluorescently stained CBA lymph node cells were analysed using FACS-II (Becton-Dickinson) as described previously^{30,31}, but with the argon ion laser set at 300 mW, 488 nm and the photomultiplier tube (EMI 9524A) at 750 V with a fluorescence gain of 8 and a light scatter gain of 4. Data for each individual mouse were derived from the analysis of 5×10^4 lymph node cells. The fluorescein isothiocyanate-coupled rabbit anti-mouse immunoglobulin (obtained from Nordic) was selected for its binding to viable cells expressing immunoglobulin at the cell surface while producing negligible staining of other cells³⁰. The values in the table were calculated after subtracting the count for immunoglobulin-positive cells and represent the mean $\pm$ s.e. calculated from seven determinations. These results have been repeated in two subsequent experiments.

Our data are consistent with recent studies of Loveland and McKenzie^{20,21} who have clearly shown that Lyt 1⁺ T cells alone (depleted of Lyt 2⁺ T cells) were sufficient to restore immune responses generated against H-2 and non-H-2 alloantigens in ATXBM (adult thymectomized, irradiated and bone-marrow reconstituted) mice. Our data also corroborate their view²¹ that Lyt 1⁺ cells have a central role in allograft responses. In either case, vitamin A acetate seems to be unique in that it enhances immune responses against an allograft when given by mouth and the mechanism of this action will require much further attention.

In summary, these findings support the hypothesis¹⁷ that the anti-cancer action of vitamin A acetate^{17,22} is mediated through an immunological process and we have shown that it acts by increasing the representation of the Lyt 1⁺2⁻ phenotype in the T-cell population.

We thank Dr G. L. Asherson for the gift of the F(ab')₂ rabbit anti-mouse immunoglobulin and Prof. I. F. C. McKenzie for the monoclonal antibodies. We thank Drs E. Simpson and B. E. Loveland for discussions and Mrs Pat McFarlane for preparation of the manuscript. This work was supported by the CRC and MRC.

Received 1 December 1982; accepted 21 January 1983.

- Moore, T. in *The Vitamins* (eds Sebrell, W. H. & Harris, R. S.) 245-266 (Academic, New York, 1967).
- Goodman, D. S. *Fedn Proc.* **38**, 2501-2503 (1979).
- Bollag, W. in *Retinoids, Advances in Basic Research and Therapy* (eds Orfanos, C. E. et al.) 5-11 (Springer, Berlin, 1981).
- Sporn, M. B. in *Retinoids, Advances in Basic Research and Therapy* (eds Orfanos, C. E. et al.) 73-76 (Springer, Berlin, 1981).
- Pawson, B. A. *Ann. N. Y. Acad. Sci.* **359**, 1-8 (1981).
- Dresser, D. W. *Nature* **217**, 527-529 (1968).
- Floersheim, G. L. & Bollag, W. *Transplantation* **15**, 564-567 (1972).
- Jurin, M. & Tannock, I. F. *Immunology* **23**, 283-287 (1972).
- Cohen, B. E. & Cohen, I. K. *J. Immun.* **111**, 1376-1380 (1973).
- Blalock, J. E. & Gifford, G. E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5382-5386 (1977).
- Dennert, G. & Lotan, R. *Eur. J. Immun.* **8**, 23-29 (1978).
- Dennert, G., Crowley, C., Kouba, J. & Lotan, R. *J. natn. Cancer Inst.* **62**, 89-94 (1979).
- Goldfarb, R. H. & Herberman, R. B. *J. Immun.* **126**, 2129-2135 (1981).
- Abb, J., Abb, H. & Deinhardt, F. *Immunopharmacology* **4**, 303-310 (1982).
- Sporn, M. B. & Newton, D. L. *Fedn Proc.* **38**, 2528-2534 (1979).
- Bollag, W. & Matter, A. *Ann. N. Y. Acad. Sci.* **359**, 9-23 (1981).
- Medawar, P. B. & Hunt, R. *Immunology* **42**, 349-353 (1981).
- Medawar, P. B., Hunt, R. & Mertin, J. *Proc. R. Soc. Lond. B* **206**, 265-280 (1979).
- Malikovsky, M. et al. *J. Immun.* **130**, 785-790 (1983).
- Loveland, B. E., Hogarth, P. M., Ceredig, R. H. & McKenzie, I. F. C. *J. exp. Med.* **153**, 1044-1057 (1981).
- Loveland, B. E. & McKenzie, I. F. C. *Transplantation* **33**, 217-221 (1982).
- Moon, R. C., Grubbs, C. J. & Sporn, M. B. *Cancer Res.* **36**, 2626-2630 (1976).
- Malikovsky, M., Asherson, C. L., Stockinger, B. & Watkins, M. C. *Nature* **300**, 652-655 (1982).
- Hughes, W. L. et al. *Fedn Proc.* **23**, 640-648 (1964).
- Julius, M. H., Simpson, E. & Herzenberg, L. A. *Eur. J. Immun.* **3**, 645-649 (1973).
- Mage, M. G., McHugh, L. L. & Rothstein, T. L. *J. Immun. Meth.* **15**, 47-56 (1977).
- Mage, M. et al. *Eur. J. Immun.* **11**, 228-235 (1981).
- Zembala, M. A., Asherson, G. L., James, B. M. B., Stein, V. E. & Watkins, M. C. *J. Immun.* **129**, 1823-1829 (1982).
- Gorini, G., Medgyesi, G. A. & Doria, G. *J. Immun.* **103**, 1132-1142 (1969).
- Abelsira, O., Edwards, A. & Simpson, E. *Eur. J. Immun.* **11**, 275-281 (1981).
- Simon, M. M. et al. *Eur. J. Immun.* **11**, 246-250 (1981).

Transfer of a cloned immunoglobulin light-chain gene to mutant hybridoma cells restores specific antibody production

Atsuo Ochi, Robert G. Hawley,
Marc J. Shulman* & Nobumichi Hozumi

Ontario Cancer Institute and Department of Medical Biophysics,
University of Toronto, 500 Sherbourne Street, Toronto,
Canada M4X 1K9

* Rheumatic Disease Unit, Wellesley Hospital, and
Department of Medical Biophysics, University of Toronto,
Toronto, Canada M4Y 1J3

The expression of immunoglobulin (Ig) genes is regulated at several levels. For example, although κ -chain production requires a DNA rearrangement that juxtaposes variable and joining segments¹, this rearrangement is not sufficient for κ -chain gene expression; that is, some cell types do not permit immunoglobulin production²⁻⁴. The mechanisms responsible for the regulation of the expression of rearranged immunoglobulin genes are poorly understood. The technique of modifying cloned genes *in vitro* and transferring the modified genes to cells in culture provides a tool for identifying the structural features required for gene expression. To analyse immunoglobulin genes in this manner, however, it is first necessary to use, as recipients, cells that normally permit immunoglobulin production. We report here that a cloned κ -chain gene is expressed in immunoglobulin-producing hybridoma cells. Furthermore, the product of the transferred κ -chain gene is capable of restoring specific antibody production to the transformed cells.

The hybridoma Sp603 produces IgM(κ) specific for the hapten 2,4,6-trinitrophenyl (TNP)⁵. The rearranged gene encoding the TNP-specific κ chain (κ_{TNP}) has been cloned⁶. As recipient cells, we used the mutant cell line igk-14 which was derived from Sp603 and does not produce the κ_{TNP} chain. As shown in Fig. 3, the κ_{TNP} gene is apparently deleted from the igk-14 cell line. Because the igk-14 cells still produce the TNP-specific μ heavy chain, it would be expected that the expression of the κ_{TNP} gene in these cells would restore the production of TNP-specific IgM. To select for cells that have taken up the κ_{TNP} gene, the cloned κ_{TNP} gene, designated *Tk1*, was inserted into the plasmid pSV2-*neo*, which contains a bacterial phosphotransferase gene (*neo*) and confers resistance to the aminoglycoside antibiotic G418 (ref. 7). In the present study, we have used two recombinant plasmids: in the plasmid pT-*Tk1*, *Tk1* was inserted so that the direction of transcription would be in the same direction as transcription initiated from the SV40 early promoter; in pR-*Tk1*, the orientation of the gene is reversed (Fig. 1).

To transfer the DNA, bacteria harbouring the pT-*Tk1* and pR-*Tk1* plasmids were converted to protoplasts and fused with igk-14, after the method of Schaffner⁸. G418-resistant transformants of igk-14 were then selected as described in the legend to Table 1. The frequency of G418-resistant cells ranged from 5×10^{-4} to 10^{-5} per input igk-14 cell, and was comparable for the pT-*Tk1*, pR-*Tk1* and pSV2-*neo* plasmids. The resistant cells were then tested for the ability to make TNP-specific plaques. By this criterion about 10% of the G418-resistant transformants from pT-*Tk1* and about 30% of the transformants from pR-*Tk1* made TNP-specific IgM. By contrast, of 30 G418-resistant transformants from the pSV2-*neo* vector alone, none made TNP-specific plaques.

Representative transformants were selected, cloned by limiting dilution and studied further. The production of the κ_{TNP} chain has been assayed by TNP-specific haemagglutination and

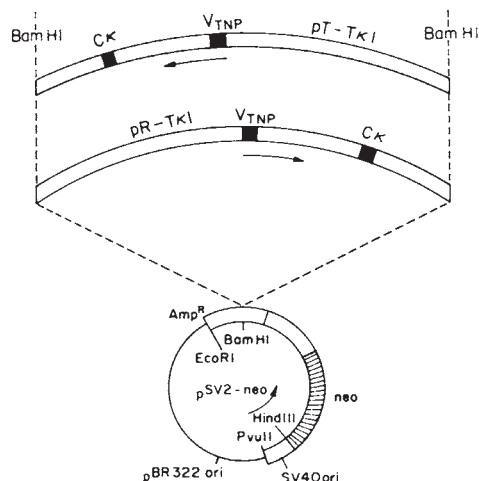


Fig. 1 Structure of *Tk1* transducing plasmids. The *Tk1* fragment (9.6 kb)⁶ was inserted into the *Bam*HI site of pSV2-*neo* (donated by P. Berg)⁷ and transfected into *Escherichia coli* cells (C600) to obtain the two types of recombinant, pT-*Tk1* (*Tk1* inserted in tandem with the SV40 early promoter) and pR-*Tk1* (*Tk1* orientation reversed with respect to the SV40 early promoter). The directions of transcription of the κ_{TNP} gene and SV40 early region are indicated by arrows.

plaque formation (Table 1). Synthesis of the κ_{TNP} chain has also been measured by SDS-polyacrylamide gel electrophoresis (PAGE) (Fig. 2). The recipient cells igk-14 still produce the κ chain from the myeloma X63-Ag8 used to generate the parental Sp603 hybridoma, and this myeloma κ chain migrates more slowly than the κ_{TNP} protein. Low level production of the κ_{TNP} chain by R11L3 can be detected both as low titre haemagglutination and as a weak band on SDS-PAGE. In general, the intensities of the bands corresponding to the κ_{TNP} chain are roughly proportional to the TNP-specific haemagglutination titres.

To analyse the κ_{TNP} genes in the transformed cell lines, transformant DNA was digested with the restriction endonuclease *Bam*HI, fractionated by agarose gel electrophoresis and transferred to nitrocellulose. The blot was then probed with a cDNA clone of the κ constant-region gene segment (Fig. 3). Previous work has shown that this probe detects three κ -chain genes in DNA from the Sp603 hybridoma—the bands at 5.9 and 5.4 kilobases (kb) correspond to κ -chain genes donated by the myeloma parent; the band at 9.6 kb represents the κ_{TNP} gene⁶, and this band is not observed in the case of igk-14. For all the G418-resistant transformants tested, including the T1L2 cells which do not make TNP-specific IgM, a band at 9.6 kb was observed with this probe. Results from DNA blot analysis of high molecular weight transformant DNA before and after digestion with the restriction endonuclease *Eco*RI (which cleaves the recombinant plasmids once) suggest that the transferred sequences are integrated into cellular DNA as tandem oligomers (data not shown). These results are consistent with those obtained by other investigators using the pSV2 transfer vectors^{7,9,10}.

We can estimate the copy number of the κ_{TNP} gene in the transformants by comparing the intensity of the band corresponding to the κ_{TNP} gene in DNA from the transformants with the intensity of the band corresponding to the κ_{TNP} gene in the DNA of Sp603 which apparently contains one copy of the κ_{TNP} gene per cell⁶. By this criterion, the transformants T3L2 and R31L4 have multiple copies of the κ_{TNP} gene per cell, while T1L2 and R11L3 contain about 1 copy per cell. It is interesting that the transformant R31L4 makes more κ_{TNP} chain than does T3L2, although T3L2 has more copies of the κ_{TNP} gene. Furthermore, R31L4 makes about 10-fold less κ_{TNP} chain per gene copy than does the wild-type hybridoma, although it should be pointed out that we do not know if all copies of the κ_{TNP} gene in R31L4 function equally efficiently. This variability in gene

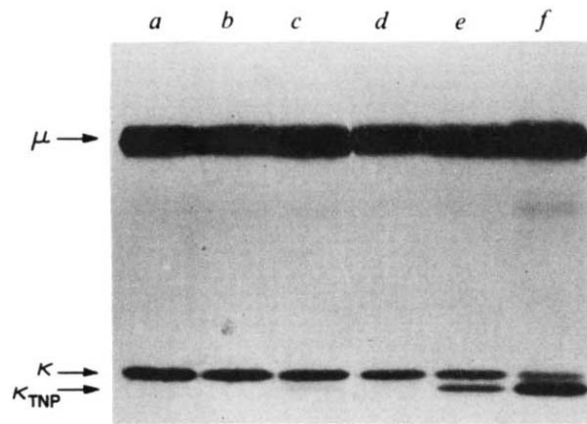


Fig. 2 Identification of κ_{TNP} -chain production in hybridoma cell lines and pT-*Tk1* and pR-*Tk1* transformants. Lane a, igk-14; lane b, T1L2; lane c, T3L2; lane d, R11L3; lane e, R31L4; lane f, Sp603. Secreted immunoglobulin was radiolabelled by incubating cells for 18 h in leucine-free Dulbecco's modified Eagle's medium containing ¹⁴C-leucine (5 μ Ci ml⁻¹) and dialysed fetal calf serum (5%). Immunoglobulin was reduced and culture supernatants analysed by SDS-PAGE as described previously⁵.

Table 1 Secretion of TNP-specific IgM by G418-resistant transformants

		Haemagglutination titre	TNP plaque formation
Cell lines			
Vector	Sp603	1/1,280	+
	igk-14	<1/2	—
	Transformants		
	T1L2	<1/2	—
	T3L2	1/160	+
	T4L1	<1/2	—
	T12e	<1/2	—
	T17L1	1/20	—
	T19L1	<1/2	—
	T21L2	<1/2	—
pR- <i>Tk1</i>	R2L6	1/1,280	+
	R9L3	<1/2	—
	R10L18	<1/2	—
	R11L3	1/20	—
	R20L1	<1/2	—
	R22L1	1/640	+
	R31L4	1/320	+

The igk-14 cells were fused with bacterial protoplasts containing either the pT-*Tk1* or pR-*Tk1* plasmids. The methods for increasing the plasmid copy number and making protoplasts have been described elsewhere¹⁴. About 10¹⁰ protoplasts were collected by centrifugation. 10⁷ igk-14 cells, grown as described previously⁵, were washed and resuspended in Gibco H21 medium, and centrifuged onto the protoplast pellet. The protoplasts and igk-14 cells were then fused with polyethylene glycol 1000 (PEG), as described by Taniguchi and Miller for the production of T-cell hybridomas¹⁵. Briefly, after centrifugation, the cell-protoplast pellet was gently resuspended in 2 ml of H21 medium containing 40% PEG and 10% dimethyl sulphoxide. After 15 s, the mixture was added to an equal volume of H21 medium containing 50% PEG and mixing was continued for a further 15 s. The cells were then diluted to 50 ml with H21 medium containing 20% fetal calf serum, washed once in the same medium and plated in 24-well culture dishes at a cell density of 10⁵ cells per well. After 24 h, the fused cells were selected in H21 medium containing 1 mg ml⁻¹ of G418 (a gift of Schering Corporation). The transformants resistant to G418 were screened for the production of TNP-specific IgM by testing for TNP-specific plaque formation. Representative transformants were then cloned by limiting dilution. Culture supernatants were tested for haemagglutination of TNP-coupled sheep red cells (TNP-SRC) as described elsewhere⁵. The sensitivity of the assay was enhanced by including monoclonal rat anti- μ antibody C2-23, as described elsewhere (M. Potash *et al.*, in preparation). Plaques were done on ~10⁴ cells, also as described previously⁵.

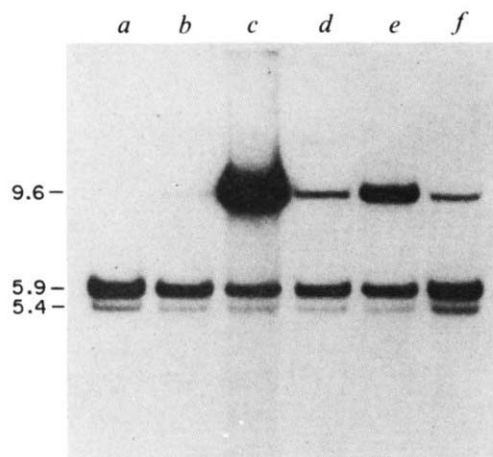


Fig. 3 Blot hybridization of DNAs from hybridoma cell lines and pT-*Tk1* and pR-*Tk1* transformants. Lane a, igk-14; lane b, T1L2; lane c, T3L2; lane d, R11L3; lane e, R31L4; lane f, Sp603. *Bam*HI-digested DNA samples (20 µg) were electrophoresed through a 1% agarose gel at 2 V cm⁻¹ for 40 h. After transfer to nitrocellulose¹⁶, the blot was hybridized with a ³²P-labelled cDNA clone of the κ constant-region gene segment (pL21-5, provided by R. Wall) by a method described elsewhere⁶.

expression raises the question of whether all the regulatory elements of the normal κ_{TNP} gene are present or functioning on the cloned fragment. The messenger RNA cap sites have been identified for several immunoglobulin κ -chain genes¹¹. In all cases, the cap site seems to be within 30 base pairs (bp) of the initiation codon. As the *Tk1* fragment has 5 kb of DNA upstream of the initiation codon, it is likely that the DNA fragment carries the κ_{TNP} gene promoter. Similarly, the poly(A) addition site is estimated to lie 211 bp downstream of the constant-region gene segment¹², which is well within the 1.2 kb of downstream flanking DNA. Experiments are in progress to determine the molecular basis for the differences in the level of expression of the κ_{TNP} gene in the various transformants.

Falkner and Zachau have studied the transient expression of mouse immunoglobulin κ -chain genes in African green monkey cells (CV1), HeLa cells and mouse L cells¹³. In the case of the monkey CV1 cells and the HeLa cells, the κ -chain gene was expressed only when transcription of the gene was placed under the direct control of an SV40 promoter after removal of the putative κ -chain gene promoter region. In the case of the mouse L cells, the κ -chain gene was not expressed. There are many structural differences between the vectors used by Falkner and Zachau and those used here. On the other hand, differences in the cellular environments might underlie the different requirements for κ -chain gene expression. It should be possible to resolve these differences by using the gene transfer system described here on cells of different types.

This work was supported by grants from the MRC of Canada, the National Cancer Institute of Canada, the Arthritis Society of Canada and the Allstate Foundation. A.O. was supported by a research fellowship of the National Cancer Institute of Canada. R.G.H. was supported by a studentship of the MRC of Canada.

Note added in proof: Rice and Baltimore¹⁷ and Oi *et al.*¹⁸ have recently reported the expression of cloned κ light chain genes in transformed lymphoid cells.

Received 17 November 1982; accepted 19 January 1983.

- Seidman, J. G., Max, E. E. & Leder, P. *Nature* **280**, 370-375 (1979).
- Parslow, T. G. & Granner, D. K. *Nature* **299**, 449-451 (1982).
- Coffino, P., Knowles, B., Nathenson, S. G. & Scharff, M. D. *Nature new Biol.* **231**, 87-90 (1971).
- Ar-Rushdi, A., Tan, K. B. & Croce, C. M. *Somatic Cell Genet.* **8**, 151-161 (1982).
- Köhler, G. & Shulman, M. J. *Eur. J. Immun.* **10**, 467-476 (1980).
- Hawley, R. G., Shulman, M. J., Murialdo, H., Gibson, D. M. & Hozumi, N. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7425-7429 (1982).
- Southern, P. J. & Berg, P. *J. molec. appl. Genet.* **1**, 327-341 (1982).
- Schaffner, W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2163-2167 (1980).
- Mulligan, R. C. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2072-2076 (1981).
- Canaani, D. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5166-5170 (1982).

- Kelley, D. E., Coleclough, C. & Perry, R. P. *Cell* **29**, 681-689 (1982).
- Hamlyn, P. H., Brownlee, G. G., Cheng, C. C., Gait, M. J. & Milstein, C. *Cell* **15**, 1067-1075 (1978).
- Falkner, F. G. & Zachau, H. G. *Nature* **298**, 286-288 (1982).
- Sandri-Goldin, R. M., Goldin, A. L., Levine, M. & Glorioso, J. C. *Molec. cell. Biol.* **1**, 743-752 (1981).
- Taniguchi, M. & Miller, J. F. A. P. *J. exp. Med.* **148**, 373-382 (1978).
- Southern, E. M. *J. molec. Biol.* **97**, 503-517 (1975).
- Rice, D. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7862-7865 (1982).
- Oi, V. T., Morrison, S. L., Herzenberg, L. A. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **80**, 825-829 (1983).

Deduced amino acid sequence from the bovine oxytocin-neurophysin I precursor cDNA

H. Land*[‡], M. Grez*[‡], S. Ruppert*, H. Schmale[†], M. Rehbein[†], D. Richter[†] & G. Schütz*

* Institute of Cell and Tumor Biology, German Cancer Research Center, Im Neuenheimer Feld 280, D-6900 Heidelberg, FRG

† Institut für Physiologische Chemie, Abteilung Zellbiochemie, Universität Hamburg, UKE, 2000 Hamburg 20, FRG

The nonapeptide hormone oxytocin-like arginine-vasopressin (AVP)¹⁻⁴ is synthesized as part of a larger precursor polypeptide. The precursor also includes the neurophysin molecule with which the hormone is associated in the neurosecretory granules of the hypothalamo-pituitary tract. A protein of molecular weight (*M_r*) ~ 20,000 has been isolated from supraoptic nuclei of rat hypothalami which, after tryptic cleavage, released a neurophysin-like molecule of *M_r* ~ 10,000 and an oligopeptide related to oxytocin⁵. This result was complemented by *in vitro* translation of bovine hypothalamic mRNA⁶⁻⁸. Among the primary translation products a single polypeptide of *M_r* ~ 16,500 was shown to contain antigenic determinants recognized by specific antisera against bovine neurophysin I and oxytocin. Here we report the amino acid sequence of the bovine oxytocin-neurophysin I (OT-NpI) precursor which was derived from sequence analysis of the cloned cDNA. As is the case for the bovine arginine-vasopressin-neurophysin II (AVP-NpII) precursor⁴, the signal sequence of the OT-NpI precursor is immediately followed by the nonapeptide hormone which is connected to neurophysin I by a Gly-Lys-Arg sequence. A striking feature of the nucleic acid sequence is the 197-nucleotide long perfect homology with the AVP-NpII precursor mRNA sequence encoding the conserved middle part of neurophysins I and II.

A cDNA library of bovine hypothalamic mRNA⁴ was screened for plasmids containing the mRNA sequence for the OT-NpI precursor. Bovine neurophysins I and II show nearly 80% homology in their amino acid sequences, including a complete homology between amino acids 10 and 74, so considerable cross-hybridization was expected between the mRNA sequences for the AVP-NpII precursor and the OT-NpI precursor.

The cloned cDNA encoding the AVP-NpII precursor⁴ was therefore chosen as a hybridization probe for *in situ* colony screening⁹. Out of 5,000 recombinants, 63 clones gave a positive hybridization signal. Restriction analysis of the plasmids revealed two groups of 47 and 16 members containing different types of cDNA inserts (see Fig. 1 for examples). The first group represented cDNA sequences specific for the AVP-NpII precursor⁴. Sequence analysis of members of the second group (Fig. 1) showed that these contained OT-NpI precursor-specific sequences.

Figure 2 shows the nucleotide sequences of the cloned OT-NpI precursor cDNA and of the previously described AVP-NpII precursor cDNA⁴ plus the predicted amino acid sequences.

[‡] Present addresses: MIT Center for Cancer Research, 77 Massachusetts Avenue, Cambridge, Massachusetts 02138, USA (H.L.); University of Southern California, School of Medicine, Department of Microbiology, 2025 Zonal Avenue, Los Angeles, California 90033, USA (M.G.).

Fig. 1 Comparison of the restriction maps of AVP-NpII precursor cDNA (pVNP-II-1, *a*), and OT-NpI precursor cDNA (p0NpI-1, *b*), including the sequencing strategy for the OT-NpI precursor cDNA (*c, d*). The double-stranded cDNA was inserted into the *Pst*I site of pBR322 as described in ref. 4. The boxed regions represent the cloned cDNA inserts including the terminal poly(dG)-poly(dC) tails (open regions). The mRNA sequence is shown as a solid bar. The predicted start site of translation (Met) and the positions of the termination codons (Term) are indicated. The regions homologous to the poly(A) tails of the mRNAs are represented by the hatched areas. To allow a simple sequencing protocol for the OT-NpI precursor cDNA, its two *Pst*I fragments were cloned into the polylinker plasmid vector pUC8 (ref. 19) (*c, d*). Five independent subclones containing different lengths of the 5' part of the mRNA sequence were sequenced (*c*). No difference in the DNA sequences was found; the known amino acid sequence of oxytocin and neurophysin I were in complete agreement with the DNA sequences. The 3' part of the OT-NpI precursor mRNA sequence was derived from sequencing two independent subclones of identical size (*d*); the two sequences did not differ. The dots and arrows indicate the restriction sites used for labelling the DNA at the 3' ends with T4 DNA polymerase and the direction and extent of DNA sequence determined respectively. ■ Represents labelling of the 5'-terminal phosphate. Preparation of the labelled DNA fragments and sequencing reactions were performed as described elsewhere⁴. bp, Base pairs.

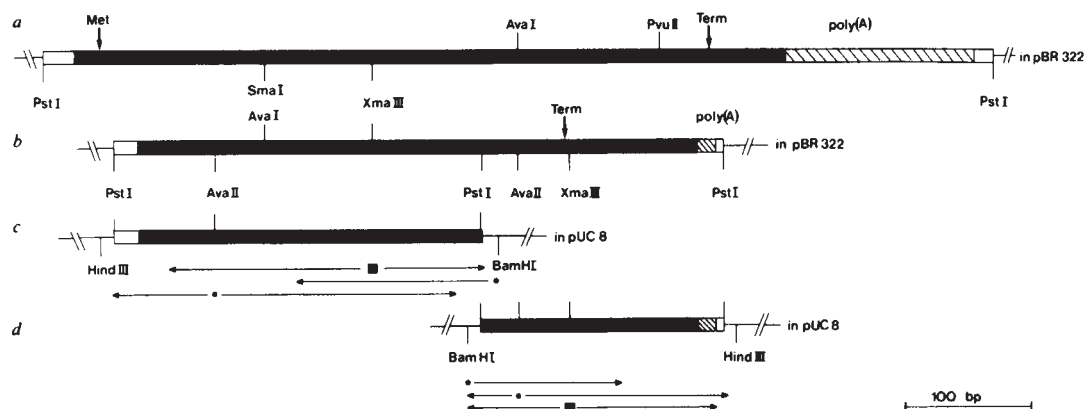


Fig. 2 Nucleotide sequence and deduced amino acid sequence of bovine OT-NpI precursor cDNA compared with the bovine AVP-NpII precursor cDNA sequence⁴. Nucleotide residues are numbered at the left side of the figure in the 5'-3' direction beginning with the first residue in the coding region for oxytocin or AVP. The nucleotides upstream from residue 1 are indicated by negative numbers. The amino acid sequences are numbered by designating as 1 the first residue (Cys) of oxytocin or AVP. The amino acids constituting the putative signal peptides are indicated by negative numbers. The parts of the nucleotide sequences encoding peptides or proteins found in the neurosecretory granula of the hypothalamo-pituitary tract are indicated by arrows. Note that the histidine at position 94 has not been found in neurophysin I. *, Identical nucleotides in homologous positions of the two cDNA sequences. A 197-nucleotide region of identity was found between positions 63 and 259.

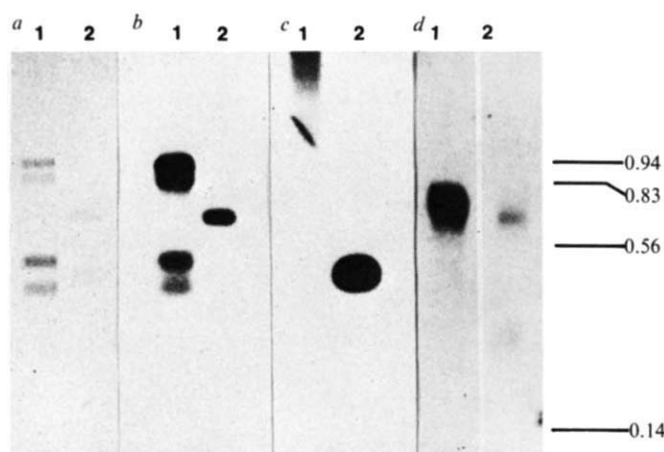


Fig. 3 Size determination of the OT-NpI precursor mRNA using an oxytocin-specific hybridization probe. The sequence comparison in Fig. 2 suggested that a probe specific for the OT-NpI precursor mRNA could be derived from nucleotides 260–435 of the cloned cDNA. The subclone containing the 3' *Pst*I fragment of the mRNA sequence (Fig. 1d) starting at position 255 was therefore tested for selective hybridization with the OT-NpI sequence. AVP-NpII precursor cDNA (Fig. 1a) was digested with *Pst*I and *Ava*I, separated on a 2% agarose gel and stained with ethidium bromide (a, lane 1). The upper band representing the 3' part of the mRNA sequence appears as a doublet because of length heterogeneity probably due to instability of the cloned poly(A) tail in *Escherichia coli* cells. OT-NpI precursor cDNA (Fig. 1b) was digested with *Pst*I (a, lane 2). In b and c the cDNA fragments were transferred to nitrocellulose and hybridized with ³²P-labelled AVP-NpII precursor cDNA (b) or with the 3'-specific subclone of the OT-NpI cDNA (c). d, Bovine hypothalamic poly(A)-containing RNA (3.0 µg per slot) was separated on a 1.5% agarose gel and transferred to a nitrocellulose filter²⁰. The size of DNA fragments used as marker is given in kilobases. In the experiment illustrated the nitrocellulose filter was first hybridized to ³²P-AVP-NpII cDNA and autoradiographed (d, lane 1), then after thorough washing the same filter was rehybridized to the ³²P-OT-NpI cDNA (d, lane 2).

Oxytocin is encoded by nucleotides 1–27. Nucleotides 37–315 (93 amino acids) correspond precisely to the amino acid sequence determined for bovine neurophysin I¹⁰. Sequence analysis of two independently isolated cDNA clones indicates that a His codon at position 106 immediately precedes the TGA termination codon. As histidine has never been found at the carboxy terminus of a known neurophysin amino acid sequence, it seems probable that the extra amino acid is removed from the primary translation product during the post-translational processing accompanying the separation of oxytocin and neurophysin I¹¹. As predicted by previous studies^{5,7,8}, the OT-NpI precursor does not carry a polypeptide moiety homologous to the glycopolypeptide of the AVP-NpII precursor^{4,5,7,8,12}.

The glycine residue at position 10 which immediately follows the amino acid sequence of oxytocin and precedes a pair of basic amino acids probably serves as a signal for the carboxy-terminal amidation of oxytocin. In all cases when the sequence of an amidated peptide has been determined from the mRNA sequence, a glycine has been found in the described position¹³. Studies on mellitin¹⁴ suggest that successive basic amino acids are not essential for the formation of the amide group, although their removal may be a prerequisite.

The cloned cDNA for the OT-NpI precursor does not contain the initiation codon for translation. It ends at residue –17 and therefore only five amino acids of the putative signal sequence could be deduced (Fig. 2). However, four of these five amino acids are identical with the homologous amino acids of the signal sequence in the AVP-NpII precursor. In the 3'-untranslated region the nucleotide sequence AATAAA, which is thought to be involved in polyadenylation, is found 19 bases upstream from the poly(A) tail.

To determine the length of the mRNA encoding the OT-NpI precursor a probe hybridizing specifically to the OT-NpI mRNA, but not to the AVP-NpII mRNA, was required. The 3' *Pst*I fragment of the cloned OT-NpI precursor cDNA was therefore subcloned (Fig. 1d). To test its specificity it was radioactively labelled and hybridized to cloned cDNA fragments carrying sequences coding for the AVP-NpII precursor (Fig. 3c, lane 1) or the OT-NpI precursor (Fig. 3c, lane 2). The probe only hybridized to itself and did not react with sequences specific for AVP-NpII precursor mRNA or with sequences common to both molecules. In contrast, the cloned cDNA for the AVP-NpII precursor does not hybridize to the 3' *Pst*I fragment of the OT-NpI precursor cDNA (Fig. 3b, lane 2), but shows strong hybridization with the larger 5' *Pst*I fragment (see also Fig. 1). The OT-NpI mRNA-specific probe was hybridized to bovine hypothalamic mRNA after electrophoresis and blotting to a nitrocellulose filter (Fig. 3d, lane 2). The RNA corresponding to OT-NpI precursor mRNA migrates with a size of about 690 nucleotides. This is smaller than the mRNA hybridizing with the AVP-NpII-specific cDNA probe (Fig. 3d, lane 1), which migrates with a length of about 750 nucleotides, as previously determined⁴. Assuming a length of 150 nucleotides for the poly(A) tail, as observed for AVP-NpII precursor mRNA⁴, it seems that the longest OT-NpI cDNA clone is missing about 100 nucleotides from the 5' end of the mRNA. The OT-NpI precursor mRNA seems less abundant than expected from the frequency of OT-NpI and AVP-NpII precursor cDNA clones. This may be related to the fact that the OT-NpI precursor mRNA can only react with a portion of the probe.

Comparison of the cDNA sequences of the OT-NpI and AVP-NpII precursors reveals a striking 197-nucleotide long homology. This region codes precisely for the highly conserved central portion of the neurophysins. An analysis of oxytocin, AVP and their related neurophysins in different species shows that the hormones emerged by gene duplication from a common ancestor about 450 Myr ago¹⁵. It is extremely unlikely that conservation of the protosequence could account for such a region of sequence identity in highly homologous, but distinct, mRNAs. It is also unlikely that this region originates from the same exon through alternative splicing pathways because in some species (pig and rat) differences in amino acid sequences are found in the conserved part of the neurophysins¹⁰. The most likely explanation for the origin of this sequence homology therefore seems to be a recent gene conversion event^{16–18}. Analysis of the genomic DNAs encoding the two mRNAs should provide further insight into the evolution of this sequence homology.

The work was supported by Deutsche Forschungsgemeinschaft Ri 192/16-3 (D.R.), Schu 51/4 (G.S.) and the Fonds der Chemischen Industrie (G.S.). We thank Dr Richard Ivell for help in preparing hypothalamic mRNA and stimulating discussions, and Mrs U. Joa for secretarial assistance.

Received 5 January; accepted 26 January 1983.

1. Brownstein, M. J. & Gainer, H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4046–4049 (1977).
2. Brownstein, M. J., Russell, J. T. & Gainer, H. *Science* **207**, 373–378 (1980).
3. Schmale, H. & Richter, D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 766–769 (1981).
4. Land, H., Schütz, G., Schmale, H. & Richter, D. *Nature* **295**, 299–303 (1982).
5. Russell, J. T., Brownstein, M. J. & Gainer, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6086–6090 (1979).
6. Schmale, H., Leipold, B. & Richter, D. *FEBS Lett.* **108**, 311–316 (1979).
7. Schmale, H. & Richter, D. *FEBS Lett.* **121**, 358–362 (1980).
8. Richter, D., Schmale, H., Ivell, R. & Kalthoff, H. in *Hormonally Active Brain Peptides* (eds McKerns, K. W. & Pantic, V.) 581–598 (Plenum, New York, 1982).
9. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961–3965 (1975).
10. Chauvet, M.-T., Chauvet, J. & Acher, R. *Biochem. biophys. Res. Commun.* **103**, 595–603 (1981).
11. Eipper, B. A. & Mains, R. E. *J. biol. Chem.* **256**, 5689–5695 (1981).
12. Ivell, R., Schmale, H. & Richter, D. *Biochem. biophys. Res. Commun.* **102**, 1230–1236 (1981).
13. Richter, D. & Schmale, H. in *Regulation of Gene Expression by Hormones* (ed. McKerns, K. W.) (Plenum, New York, in the press).
14. Suchanek, G. & Kreil, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 975–978 (1977).
15. Acher, R. *Proc. R. Soc.* **210**, 21–43 (1980).
16. Baltimore, D. *Cell* **24**, 592–594 (1981).
17. Egel, R. *Nature* **290**, 191–192 (1981).
18. Petes, T. & Fink, G. R. *Nature* **300**, 216–217 (1982).
19. Vieira, J. & Messing, J. *Gene* **19**, 259–268 (1982).
20. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).

The vasopressin receptor antagonist dTyr(Me)AVP does not prevent stress-induced ACTH and corticosterone release

Pierre Mormède

Laboratoire de Neurobiologie du Comportement INRA,
Université de Bordeaux II, 146 rue Léo Saignat,
33076 Bordeaux Cedex, France

Most of the experimental evidence for a role of arginine-vasopressin (AVP) in adrenocorticotrophic hormone (ACTH) release comes from *in vitro* studies¹. The multimolecular nature of the hypothalamic factor responsible for corticotropin (CRF) release has long been recognized², but the importance of AVP as a cofactor is controversial^{3,4}. The recently characterized 41-residue peptide^{5,6} fulfills the criteria for a physiological role in ACTH release⁷⁻⁹ and it is potentiated *in vitro* by AVP^{10,11}. *In vivo*, AVP is able to stimulate ACTH secretion¹², and Brattleboro rats, deficient in AVP, show a reduced activity of the hypothalamo-hypophysial-adrenocortical system (HHCS) (see ref. 13 for references). Direct evidence for involvement of AVP in the physiological release of ACTH is, however, still lacking. The recent development of AVP receptor antagonists¹⁴ provides the opportunity to test this hypothesis directly. I report here that pretreatment by 1-deaminopenicillamine, 2-(*O*-methyl)tyrosine arginine-vasopressin (dTyr(Me)AVP), a potent antagonist of the vasopressor^{14,15}, behavioural¹⁵ and ACTH-releasing properties of AVP¹⁶, does not modify the ACTH and corticosterone secretion induced by exposure to a novel environment, but totally inhibits the increase of ACTH and corticosterone levels induced by AVP. The results do not support the hypothesis for a physiological involvement of AVP in ACTH release.

Male Wistar rats (IFFA CREDO, 200–250 g) were exposed to a standardized novel environment which stimulates HHCS activity. This can be considered as a purely psychological activator which eliminates the more specific responses elicited by physiological stressors¹⁷. In practice, the animals were placed singly in buckets containing some sawdust litter, in a room with a background masking noise. The rats were killed by decapitation 15 or 30 min later and trunk blood was collected. Control animals were left in the home cage before decapitation.

Plasma corticosterone was measured by radiocompetitive binding¹⁸ using pregnant human (0.4–12.8 ng per tube) or rhesus monkey (0.04–1.28 ng per tube) serum as the source of transcortin, according to the steroid level, tritiated corticosterone (107 Ci mol⁻¹; CEA, Gif-sur-Yvette) as the tracer and dextran-coated charcoal as the adsorbent of free radioactivity. Plasma ACTH was measured by direct radioimmunoassay with antibodies raised against human ACTH, ¹²⁵I-labelled porcine ACTH as the tracer and charcoal as the adsorbent of free radioactivity (ACTH radioimmunoassay kit; CEA, Gif-sur-Yvette). This methodology has previously been validated by parallelism studies.

For intracerebroventricular (i.c.v.) injections, animals were surgically prepared with chronic 23-gauge stainless-steel guide cannulae stereotactically positioned¹⁹ and the peptide was administered in 2.0 µl of saline. All other treatments were subcutaneous (s.c.) and the dose, in a volume of 0.5 ml, was adjusted to the mean body weight of the group. Peptides were synthesized at the Salk Institute and were a gift of Dr J. Rivier. They were first dissolved in 50 µl of 0.1 M acetic acid and then diluted in 0.9% NaCl. The logarithmically transformed data were subjected to either a one- or two-way analysis of variance.

Exposure to a novel environment induced a marked increase of corticosterone levels from 8.3 (5.8–11.8, mean ± s.e.m.) to 233 (193–281) ng ml⁻¹ 30 min later. Neither basal nor post-

stress level was modified by s.c. anti-AVP (A-AVP; 1–25 µg per kg) administered immediately before the test (data not shown). Similarly, i.c.v. A-AVP did not change post-stress levels (controls: 360 (349–372) ng corticosterone per ml; 10 ng A-AVP per rat: 382 (363–403) ng corticosterone per ml; 100 ng A-AVP per rat: 375 (365–386) ng corticosterone per ml; 8 rats per group).

In a second experiment, the effects of higher s.c. doses were compared on stress- and AVP-induced corticosterone release, to ascertain the biological activity of the antagonist. A-AVP (s.c.) in doses up to 500 µg per kg was unable to modify either basal or stress-increased levels of corticosterone (Fig. 1). However, the corticosterone release [14.6 (12.6–16.9) to 187 (151–233) ng ml⁻¹] induced by s.c. AVP (5 µg per kg) was completely inhibited by the lowest dose of A-AVP (20 µg per kg, s.c.).

In view of the possible dissociation between ACTH and corticosterone release²⁰, ACTH was also measured 15 and 30 min after stress or AVP injection, A-AVP (100 µg per kg) or saline being administered s.c. 15 min before the test. Basically the same results were obtained: in the novel environment, A-AVP pretreatment did not alter ACTH release [851 (729–995) and 696 (697–799) pg ml⁻¹ for control and A-AVP respectively, 12 rats per group, $F(1, 20) = 0.97$], although the ACTH response to AVP was strongly attenuated by A-AVP pretreatment [from 348 (287–420) to 86 (67–111) pg ml⁻¹, 12 rats per group, $F(1, 20) = 17.91$, $P < 0.001$]. Again A-AVP did not change corticosterone response to the novel environment ($F(1, 20) = 0.75$) but reduced the AVP-induced stimulation ($F(1, 20) = 11.07$; $P < 0.01$). These results show that the whole HHCS response to psychological stress was insensitive to AVP receptor blockade.

The CRF activity of the 41-residue peptide isolated from ovine hypothalamus⁵ is potentiated several-fold by vasopressin *in vitro*¹⁰. The blockade of AVP receptors *in vivo* was therefore expected to reduce stress-induced ACTH release. The vasopressin analogue dTyr(Me)AVP has been shown to be essentially devoid of agonistic activity, apart from a very low antidiuretic activity¹⁴. Very large doses are required for an effect on conditioned behaviour²¹. We were also unable to show any change in basal corticosterone levels due to A-AVP, even

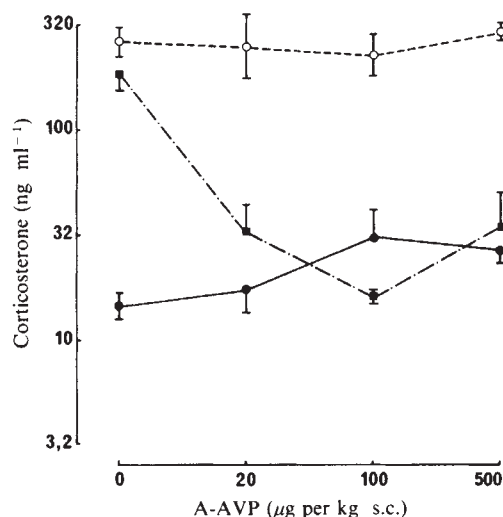


Fig. 1 Dose-response effects of s.c. A-AVP on basal levels, stress- and AVP-stimulated plasma corticosterone levels. Rats were injected s.c. with either 0.5 ml of saline or A-AVP and 15 min later were injected s.c. with either 0.5 ml of saline (base) or AVP (5 µg per kg), or were submitted to the novel environment test. All animals were killed 30 min later. A-AVP did not change either basal or stress-stimulated levels ($F(3, 16) = 2.62$ and 0.21 respectively, five rats per group), but reduced the AVP-stimulated levels ($F(3, 16) = 14.85$, $P < 0.001$, $n = 6, 6, 4$, four rats for 0, 20, 100, 500 µg per kg of A-AVP respectively). ○, Stress; ■, AVP (5 µg per kg, s.c.); ●, base.

though high doses were used. However, the antagonist is powerful in blocking the vasopressor^{14,15}, behavioural¹⁵ and ACTH-releasing¹⁶ properties of AVP. For example, the vasopressor effect of AVP was shown to be antagonized by previous administration of A-AVP in a dose ratio with AVP of 5:1 (ref. 15). A similar ratio was found here for the HHCS-activating effects. In fact, a correlation between vasopressor and CRF-like properties of AVP analogues has been found¹⁶, suggesting that similar receptors are involved in both cases. The inability of the antagonist to change the ACTH response to the psychological stress of exposure to a novel environment

Received 30 November 1982; accepted 8 February 1983.

1. Baertschi, A. J. & Bény, J.-L. *Frontiers Horm. Res.* **10**, 126-140 (1982).
2. Pearlmutter, A. F., Rapino, E. & Saffran, M. *Endocrinology* **97**, 1336-1338 (1975).
3. Gillies, G. & Lowry, P. *Nature* **278**, 463-464 (1979).
4. Pearlmutter, A. F., Dokas, L., Kong, A., Miller, R. & Saffran, M. *Nature* **283**, 697-698 (1980).
5. Vale, W., Spiess, J., Rivier, C. & Rivier, J. *Science* **213**, 1394-1397 (1981).
6. Spiess, J., Rivier, J., Rivier, C. & Vale, W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6517-6521 (1981).
7. Rivier, C., Brownstein, M., Spiess, J., Rivier, J. & Vale, W. *Endocrinology* **110**, 272-278 (1982).
8. Szeirias-Diaz, J. *et al. Life Sci.* **31**, 429-435 (1982).
9. Hook, V. Y. H., Heisler, S., Sabol, S. L. & Axelrod, J. *Biochem. biophys. Res. Commun.* **106**, 1364-1371 (1982).
10. Gillies, G. E., Linton, E. A. & Lowry, P. J. *Nature* **299**, 355-357 (1982).

can be explained by the non-involvement of hypophysial AVP receptors in ACTH release. The release of AVP under stress is not firmly established. Although it has been shown to occur under special circumstances, such as in fever²², it does not seem to be a general stress response²³. On the other hand, the hypophysial AVP receptor involved in CRF potentiation could be insensitive to the antagonistic properties of dTyr(Me)AVP.

I thank Rose-Marie Bluthé and Yvette Langlois for technical assistance and Dr Simon Jarman for manuscript correction. This work was supported by Institut National de la Santé et de la Recherche Médicale contrat libre 826 013.

11. Turkelson, C. M. *et al. Peptides* **1**, 111-113 (1982).
12. Yasuda, N., Greer, M. A., Greer, S. E. & Pantoni, P. *Endocrinology* **103**, 906-911 (1978).
13. Conte-Devolx, B. *et al. Endocrinology* **110**, 2097-2100 (1982).
14. Bankowski, K., Manning, M., Haldar, J. & Sawyer, W. H. *J. med. Chem.* **21**, 850-853 (1978).
15. Le Moal, M. *et al. Nature* **291**, 491-493 (1981).
16. Aizawa, T., Yasuda, N., Greer, M. A. & Sawyer, W. H. *Endocrinology* **110**, 98-104 (1982).
17. Mason, J. W. *J. psychiat. Res.* **8**, 323-333 (1971).
18. Murphy, B. E. P. *J. clin. Endocr. Metab.* **27**, 973-990 (1967).
19. Britton, D. R., Koob, G. F., Rivier, J. & Vale, W. *Life Sci.* **31**, 363-367 (1982).
20. Wood, C. E., Shinsako, J., Keil, L. C., Ramsay, D. J. & Dallman, M. F. *Endocrinology* **110**, 1416-1421 (1982).
21. Koob, G. F. *et al. Regulat. Peptides* **2**, 153-163 (1981).
22. Kasting, N. W., Veale, W. L. & Cooper, K. E. *Neurosci. Biobehav. Rev.* **6**, 215-222 (1982).
23. Husain, M. K., Manger, W. M., Rock, T. W., Weiss, R. J. & Frantz, A.-G. *Endocrinology* **104**, 641-644 (1979).

Genetic evidence that a Y-linked gene in man is homologous to a gene on the X chromosome

P. Goodfellow*, G. Banting*, D. Sheer*,
H. H. Ropers†, A. Caine‡, M. A. Ferguson-Smith‡,
S. Povey§ & R. Voss||

* Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

† Institute für Humangenetik und Anthropologie, Albert-Ludwigs-Universität, Freiburg, FRG

‡ Duncan Guthrie Institute of Medical Genetics, University of Glasgow, Glasgow G3 4SJ, UK

§ MRC Human Biochemical Genetics Unit, Galton Laboratories, University of London, London NW1 2HE, UK

|| Department of Human Genetics, Hadassah Medical Centre, Ein Karem, Jerusalem, Israel

The mammalian sex chromosomes are thought to be related to each other by sharing a common origin. That is, the X and Y chromosomes originally evolved from a pair of chromosomes that only differed at the locus determining sexual differentiation^{1,2}. For example, this evolutionary relationship is reflected during meiosis³ in chromosomal pairing between the tip of the human X chromosome short arm and the Y chromosome which presumably implies sequence homology⁴. However, compelling genetic evidence for functional homology between the mammalian X and Y chromosome is lacking. We describe here the localization of a gene to the tip of the short arm of the human X chromosome and evidence for a related gene on the Y chromosome.

The monoclonal antibody 12E7 was raised by Levy *et al.* against human leukaemic T cells⁵. The antigenic determinant detected by this antibody is carried on a 30,000-molecular weight protein which, although present in larger amounts on thymocytes, is expressed on all human tissues so far tested with the possible exception of spermatozoa^{6,7}. The 12E7 antigenic determinant is not found on the cell surface of mouse, rat or hamster cells and this human-rodent polymorphism has been exploited to study the genetic control of expression of 12E7 antigen⁶. In human-rodent hybrids the expression of the 12E7 antigen correlates with the presence of the human X chromosome. In addition, the antigen is expressed on the surface of hybrid cells in which an X chromosome is the only detectable human genetic contribution (ref. 6. and Fig. 1, MOG13.9). This result demonstrates that an X-linked gene, provisionally desig-

nated *MIC2* (ref. 8) controls cell-surface expression of the 12E7 antigenic determinant.

We have previously demonstrated that a polymorphism exists for the level of expression of the 12E7 antigen on red blood cells. This polymorphism is genetically related to the Xg blood group polymorphism (see ref. 9 for a review of the Xg system) and shows complex sex limitation in its expression¹⁰. The polymorphism, unlike the expression of the 12E7 antigen, is restricted to red blood cells. The Xg blood group locus has been assigned to the tip of the short arm of the X chromosome (Xp22.3) by consideration of individuals with X-chromosome deletions^{9,11}. To investigate the relationship between the X-linked *MIC2* locus, the locus controlling the 12E7 quantitative polymorphism and the Xg locus, we investigated the localization of *MIC2* on the human X chromosome. A series of human-mouse somatic cell hybrids containing different fragments of the X chromosome were investigated for expression of the 12E7 antigen. The results are summarized in Fig. 1. By this deletion analysis, *MIC2* maps to the short arm of the X chromosome. A more precise map position was obtained by studying hybrids derived from fibroblasts of two unrelated male individuals (AMCl (ref. 11) and 445 (ref. 12)) who have deletions of the tip of the X-chromosome short arm (that is, lack Xp22.3-pter). Both males inherited an X-Y translocation chromosome (Xqter-Xp22.3:Yq1.1-Yqter) from their

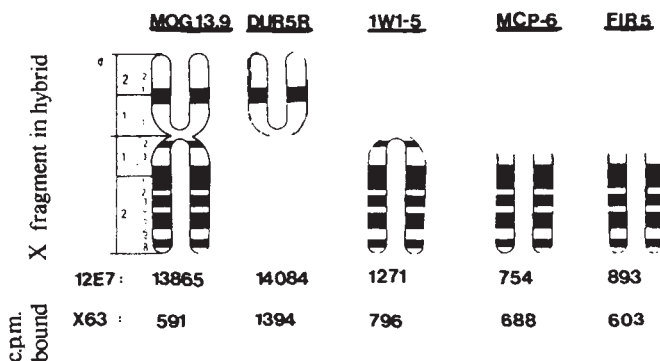


Fig. 1 Deletion mapping for localizing the human gene *MIC2* on the X chromosome. Details of the radioimmunoassay are given in the legend to Table 1. The negative control (X63) was the irrelevant myeloma antibody P3.X63.Ag8³⁶. The growth conditions, origin and characterization of the hybrids are described in refs 37 (MOG13.9), 38 (DUR5R), 39 (IWI-5), 40 (MCP-6) and 41 (FIR5). The short arm of the human X chromosome in DUR5R was identified by staining for the associated chromosome 15 centromere using the distamycin and DAPI method⁴². Although earlier transfer generations of DUR5R contained the inactive human X chromosome³⁸, it was not present in cells analysed here.

Table 1 Localization of *MIC2* to the region Xp2.23-Xpter

	Human chromosomes present	c.p.m. bound (s.d.)	
		12E7	P3.X63.Ag8
Hybrid			
AMIR.2XI	1, 3, 4, 6, 7, 11, 13, 14, 15, 17, 18, 19, 20, 21, 22, X-Y	1,066 (35)	1,026 (69)
445 × 393	1, 3, 7, 12, 13, 18, 20, X-Y	483 (70)	402 (25)
Human cells			
AMcI	Parent of AMIR-2XI carries X-Y translocation	<u>24,542 (798)</u>	477 (69)
44562	Parent of 445 × 393 carries X-Y translocation	<u>23,753 (2453)</u>	294 (47)
Mouse cells			
1R	Mouse L cell	503 (43)	640 (46)

AMIR.2 was produced by fusing the HPRT⁻ mouse cell IR²⁵ and AMcI using polyethylene glycol²⁶ and selecting in hypoxanthine-aminopterin-thymidine medium²⁷ supplemented with ouabain²⁸. AMIR.2XI is a clone of AMIR.2. 445 × 393 was produced in a similar way by fusing 445 to the mouse L cell 393 (H.H.R., unpublished results). Characterization of the human chromosomes present was by a combination of isozyme analysis²⁹, antigen analysis³⁰ and karyotypic analysis (using G11 staining³¹, trypsin banding³² and quinacrine banding³³). The binding assay was performed as described in Goodfellow *et al.*⁶ using the method of Williams³⁴. The 12E7 antibody was used at a concentration of 20 µg ml⁻¹, the class-matched negative control P3.X63.Ag8 (X63)²³ was used at the same concentration. The values represent the average of four determinations. Arbitrarily, three times the P3.X63.Ag8 background was chosen as a positive reaction; positive values are underlined.

Table 2 Reaction of 12E7 with derivatives of the hybrid AMIR.2

	Human autosomes present	% Cells containing:		c.p.m. bound (s.d.)	
		Human X-Y	Human Y	12E7	P3.X63.Ag8
Original clone					
AMIR.2	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22	100	56	<u>7,872</u> (860)	635 (101)
Non-selected subclones					
AMIR2.I	1, 2, 3, 4, 5, 6, 7, 10, 11, 13, 14, 15, 16, 17, 18, 20, 21, 22	100	90	<u>12,763</u> (2,170)	772 (56)
AMIR2.II	1, 2, 3, 5, 6, 7, 10, 11, 13, 14, 15, 16, 17, 18, 20, 21, 22	100	80	<u>4,925</u> (649)	797 (128)
AMIR2.IV	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22	96	96	<u>11,577</u> (523)	913 (122)
AMIR2.VIII	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 13, 14, 15, 17, 18, 19, 20, 21, 22	96	96	<u>9,240</u> (1,324)	920 (115)
AMIR2.XI	1, 3, 4, 6, 7, 11, 13, 14, 15, 17, 18, 19, 20, 21, 22	95	0	690 (4)	567 (69)
AMIR2.XV	1, 2, 3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22	100	80	<u>15,901</u> (1,841)	1,345 (276)
FACS-selected mass cultures					
AMIR2N	2, 3, 4, 5, 11, 13, 14, 15, 17, 18, 19, 20, 21, 22	95	0	970 (266)	928 (87)
AMIR2.VIII N3	1, 2, 3, 5, 6, 7, 10, 11, 13, 14, 15, 17, 18, 19, 20, 21, 22	94	8	1383 (65)	1,162 (102)

The origin of AMIR.2 is described in Table 1 legend. AMIR.2I, AMIR.2II, AMIR.2IV, AMIR.2VIII, AMIR.2XI and AMIR.2XV are independent clones, isolated by low density plating, derived from AMIR.2. AMIR2N was isolated from AMIR2 by selecting for the 12E7-negative population using the FACS; similarly, AMIR2.VIII N3 was isolated from AMIR2.VIII. The conditions for the sort are described in Fig. 2 legend. Characterization of the human autosomes was as described in Table 1 legend. In those cases where both isozyme analysis and karyotyping were performed, complete concordance was found except for the presence of chromosome 8 in AMIR2.IV, which was seen in 84% of the cells; however, the isozyme marker GSR was not detected. X-Y and Y chromosome-containing cells were scored using quinacrine banding³³. The binding assay is described in Table 1 legend. Positive values are underlined.

mothers and a normal Y chromosome from their fathers. Thus, they completely lacked the region of the X chromosome contained between band Xp22.3 and Xpter. Both individuals suffer from the disease X-linked ichthyosis which is the direct result of steroid sulphatase deficiency. The *STS* locus has thus been assigned to the missing part of the short arm of the X chromosome^{12,13}. In addition, the X-Y translocation present in the cells from individual AMcI has been shown by family studies to be deleted for the *Xg* locus¹¹, while studies on the red blood cells of the family of individual 445 have proved uninformative for the *Xg* locus. Fibroblasts from both males were fused with mouse L cells to produce hybrids carrying the X-Y translocation. Hybrids that carried the X-Y translocation and no other human sex chromosomes fail to express the 12E7 antigen, confirming the assignment of *MIC2* to the short arm of the X chromosome and further localizing the gene to the terminal band Xp22.3, the same region of the X chromosome as that containing the *Xg* and *STS* loci (Table 1).

We predicted from the deletion of *MIC2* from the X-Y translocation chromosome that the fibroblasts from AMcI and 445 should both be 12E7 antigen negative. However, this was not the case (Table 1). From studies using somatic cell hybrids we have shown that the additional 12E7-controlling locus in these cells is Y-linked.

The hybrid AMIR2, produced by fusing AMcI fibroblasts with L cells, is 12E7 antigen positive (Table 2). Karyotypic analysis demonstrated the presence of the selected X-Y translocation in 92% of the population and the human Y chromosome in about 56% of the population. Clones of AMIR2 were obtained by plating cells at low density and picking independent colonies from different dishes. Of six clones analysed, five 12E7

antigen positive and one antigen negative. The only human chromosomes to show concordant segregation with 12E7 antigen expression are chromosome 2, 5 and Y. 12E7 antigen-negative mass populations were selected, using the fluorescence activated cell sorter (FACS), from both AMIR2 and the subclone AMIR2.VIII (see Table 2). AMIR2N, the antigen-negative population derived from AMIR2, lacks a Y chromosome and contains both chromosomes 2 and 5. Similarly, AMIR2.VIII N3, the antigen-negative population isolated after three rounds of selection from AMIR2.VIII, also retained chromosomes 2 and 5. Thus in both cases the absence of detectable 12E7 antigen correlates with the low percentage of cells with a Y chromosome.

Confirmation of Y-linkage of a 12E7-controlling gene was obtained from another independently derived human-mouse hybrid. Marcus *et al.* fused human peripheral blood lymphocytes with the mouse adenocarcinoma RAG. A hybrid clone was isolated that contained the human Y chromosome as its only human genetic contribution. During culture of these hybrid cells, the Y chromosome underwent a rearrangement duplicating part of the Y chromosome¹⁴. Table 3 shows that there was a correlation between the presence of the Y chromosome in subclones of the original Y-containing hybrid and expression of the 12E7 antigen. One clone, 3G7, was chosen for further study, because analysis with the FACS suggested that this clone was composed of a mixed population of cells with 30% failing to react with 12E7 (Fig. 2a). The positive and negative subpopulations were separated using the FACS (Fig. 2c, d) and karyotyped for the presence of the Y chromosome. All cells of the 12E7 antigen-positive subpopulation contained the Y chromosome whereas the 12E7 antigen-negative subpopulation

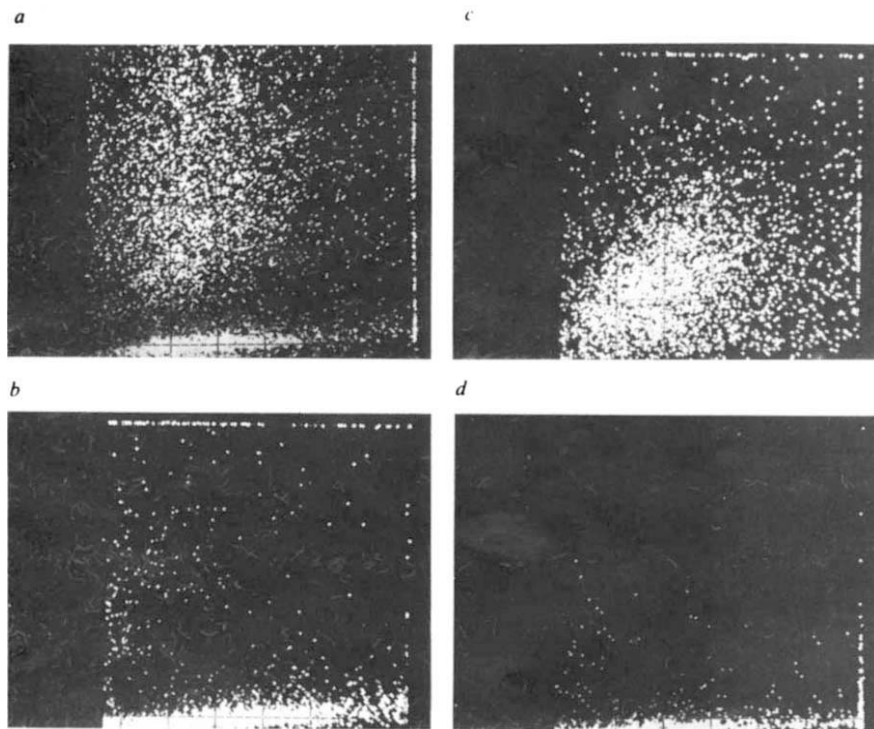


Fig. 2 Isolation of 12E7 antigen-positive and -negative subpopulations from 3G7. *a*, Staining 3G7 with 12E7; *b*, staining 3G7 with P3.X63.Ag8; *c*, staining the 12E7 antigen-positive subpopulation of 3G7 with 12E7; *d*, staining the 12E7 antigen-negative subpopulation of 3G7 with 12E7. The horizontal axis represents light scattering (which is proportional to cell size), the vertical axis represents fluorescence intensity. Each spot represents a single cell and 5,000 cells are displayed in each photograph. The flow microfluorimetry was performed as previously described³⁵. The first antibody was used at a concentration of 50 $\mu\text{g ml}^{-1}$; the second antibody (fluorescein-conjugated rabbit anti-mouse immunoglobulin; Cappel) at a dilution of 1:4. The cell numbers and volumes were scaled up for sorting the antigen-positive and -negative cells. For isolating the antigen-positive population the brightest 50% of 12E7 antigen-positive cells were collected. For isolating the negative cells the 10% dimmest cells were collected. The collected cells were grown for 2 weeks and then subjected to analysis for 12E7 antigen (*c*, *d*), karyotyping and cloning (see Table 3).

Table 3 Reaction of 12E7 with 'Y only' hybrids

Hybrid clone	Human Y chromosome	c.p.m. bound (s.d.)	
		12E7	P3.X63.Ag8
3E7	+	34,544 (1,278)	680 (47)
3A11	+	31,916 (822)	554 (12)
3F3	+	13,097 (473)	209 (23)
2F12	+	21,626 (1,097)	747 (95)
3B11	+	27,420 (1,214)	307 (37)
3G7	+	20,923 (1,432)	814 (199)
7/1	—	742 (63)	660 (92)
7/2	—	569 (17)	491 (188)
2E4	—	969 (107)	475 (42)
3G7 clones			
3G7P2	+	20,667 (665)	524 (115)
3G7P4	+	27,862 (1,617)	525 (166)
3G7N2	—	437 (61)	296 (20)
3G7N5	—	767 (32)	460 (150)
Mouse parent			
RAG	—	808 (54)	865 (31)

All the hybrids are clones, isolated by Marcus and R.V., from the hybrid described by Marcus *et al.*¹⁴. The clones of 3G7 were derived from the 12E7 antigen-positive sorted population (see Fig. 2, 3G7P2 and 3G7P4) and the 12E7 antigen-negative population (3G7N2 and 3G7N5). The human Y chromosome was recognized by quinacrine banding³³. +, One or more copies of the human Y chromosome in >50% of the cells (many cells have multiple copies of the Y chromosome, see ref. 14). —, Less than 10% of the cells contain a human Y chromosome. The binding assay is as described in Table 1 legend. Positive values are underlined.

lacked the human Y chromosome. This difference was also found in the clones derived from the antigen-positive and -negative subpopulations (Table 3). The antigen-positive hybrids contain the Y chromosome as the only detectable human genetic material. This result strongly suggests that there is a gene controlling expression of the 12E7 antigen on the Y chromosome. The finding that human-mouse hybrids containing the X-Y translocation (see Table 1), which includes Yq1:1-Yqter, are 12E7 antigen-negative, localizes the Y-linked 12E7 gene to the euchromatin region Yq1:1-Ypter.

Knowledge of the genetics and function of the human Y chromosome remains rudimentary. The meiotic pairing of the tip of the short arms of the human X and Y chromosomes suggests that the paired regions share sequence homology. At least two sequences have been demonstrated to be shared by the X and Y chromosomes^{15,16}. Karyotypic analysis of individuals suffering from sexual dysgenesis has implicated the

presence of control genes for testis determination on the Y chromosome¹⁷. Control genes for growth rates have also been assigned to the Y chromosome¹⁸. Evidence for structural genes on the Y chromosome is limited to the H-Y antigen; however, the structural gene for the serologically defined H-Y-related antigen is apparently not Y-linked although this is an area of some debate^{19,20}. The present report gives evidence for a functional gene on the Y chromosome, localizes this gene to the euchromatin region (Ypter-Yq1:1) and implies that the X and Y chromosomes still share at least one functional gene.

The 12E7 antigenic determinant is restricted to man, chimpanzees and gorillas⁷. The finding that this determinant is coded for by sequences on both the X and Y chromosomes argues for at least one exchange between the sex chromosomes in the human lineage. The 12E7 determinant is a marker for a Y-linked locus that might be a part of the ancient sex chromosome pair deriving from a period morphological and functional differentiation of the chromosomes occurred. Unlike most of the genetic information present on the original heterogametic chromosome, this locus has remained on the Y chromosome. This maintenance of Y-linkage could be random or, alternatively, this locus might be selectively retained on the Y chromosome either because of a function in sexual differentiation or because it is trapped in a region forced to retain sequence homology in order to function in meiotic pairing. In this regard it is interesting that the X-linked *MIC2* gene lies in the same region as the two genes *Xg* and *STS* known to escape X-inactivation^{9,21,22}. Preliminary experiments (P.G. *et al.*, unpublished) using somatic cell hybrids retaining inactive human X chromosomes suggest that the *MIC2* locus also escapes X-inactivation. This implies that the functional dosage of the 12E7 antigen-producing genes is the same in males and females.

The presence of a structural gene for 12E7 in Xp22.3 and on the Y chromosome also has implications for our previous work on the red blood cell polymorphism for 12E7 expression. The finding that the *Xg* locus and *MIC2* map closely together suggests that these loci may be related or even identical. Similarly, the Y-linked 12E7-controlling locus may be identical to the postulated Yg locus¹⁰. Polymorphism at the *Xg* locus and the Yg locus shows similar allele frequencies. This could be due to chance, to selection, or to recombination between the X and Y chromosomes^{23,24}.

We thank P. W. Andrews, B. O. Bengtsson, F. Benham, W. F. Bodmer, C. Carlin, R. Sanger, E. Solomon, P. Tippet, J.

Trowsdale and A. Tunncliffe for comments on this manuscript; C. Furse for typing the manuscript and R. Levy for generously donating the 12E7 antibody. Part of this work was supported by US-Israel Binational Science Foundation, grant 1962-79.

Received 30 November 1982; accepted 2 February 1983.

- Ohno, S. *A. Rev. Genet.* **3**, 495-524 (1969).
- Ferguson-Smith, M. A. *J. med. Genet.* **2**, 145-155 (1965).
- Pearson, P. C. & Bobrow, M. *Nature* **226**, 959-961 (1970).
- Rasmussen, S. W. & Holm, P. B. *Hereditas* **93**, 187-126 (1980).
- Levy, R., Dilley, J., Fox, R. I. & Warnke, R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6552-6556 (1979).
- Goodfellow, P., Banting, G., Levy, R., Povey, S. & McMichael, A. *Somatic Cell Genet.* **6**, 777-787 (1980).
- Goodfellow, P. N. *Differentiation* (in the press).
- Report of the Nomenclature Committee *Cytogenet. Cell Genet.* **32**, 221-245 (1982).
- Race, R. & Sanger, R. *Blood Groups in Man* 6th edn (Blackwell, Oxford, 1975).
- Goodfellow, P. N. & Tippet, P. *Nature* **289**, 404-405 (1981).
- Ferguson-Smith, M. A., Sanger, R., Tippet, P., Aitken, D. A. & Boyd, E. *Cytogenet. Cell Genet.* **32**, 273-274 (1982).
- Tiepolo, L. *et al. Hum. Genet.* **54**, 205-206 (1980).
- Mohandas, T., Shapiro, L. J., Sparkes, R. S. & Sparkes, M. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5779-5783 (1979).
- Marcus, M., Trantravali, R., Der, V. G., Miller, D. A. & Miller, O. J. *Nature* **262**, 63-65 (1976).
- Daigler, S. P., Wildin, R. S. & Su, T.-S. *Nature* **298**, 682-684 (1982).
- Page, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 5352-5356 (1982).
- Davis, R. J. *med. Genet.* **18**, 161-197 (1981).
- Tanner, J. M., Prader, A., Habich, H. & Ferguson-Smith, M. A. *Lancet* **ii**, 141 (1959).
- Silvers, W. K., Gasser, D. C. & Eicher, E. M. *Cell* **28**, 439-440 (1982).
- Goodfellow, P. N. & Andrews, P. W. *Nature* **295**, 11-13 (1982).
- Shapiro, L. J., Mohandas, T., Weiss, R. & Romeo, G. *Science* **204**, 1224-1226 (1979).
- Muller, C. R., Migl, B., Traupe, H. & Ropers, H. H. *Hum. Genet.* **54**, 197-199 (1980).
- Darlington, C. D., Haldane, J. B. S. & Koller, P. C. *Nature* **133**, 417-419 (1937).
- Burgoyne, P. S. *Hum. Genet.* **61**, 85-90 (1982).
- Nabholz, M., Miggiano, V. & Bodmer, W. F. *Nature* **223**, 358-363 (1969).
- Ponte-corvo, G. *Somatic Cell Genet.* **1**, 397-400 (1975).
- Szybalski, W., Szybalska, E. & Rayni, G. *Natn. Cancer Inst. Monogr.* **7**, 75-89 (1962).
- Kucherlapati, R. S., Baker, R. & Ruddle, F. H. *Cytogenet. Cell Genet.* **14**, 192-193 (1975).
- Harris, H. & Hopkinson, D. A. in *Handbook of Enzyme Electrophoresis in Human Genetics* (North-Holland, Amsterdam, 1977).
- Goodfellow, P. N. & Solomon, E. in *Monoclonal Antibodies in Clinical Medicine* (eds McMichael A. & Fabre, J.) 365-393 (Academic, London, 1982).
- Bobrow, M. & Cross, J. *Nature* **251**, 77-79 (1974).
- Seabright, M. *Chromosoma* **36**, 204-210 (1972).
- Caspersson, T., Zech, L. & Johansson, C. *Hereditas* **67**, 89-102 (1971).
- Williams, A. F. *Eur. J. Immun.* **3**, 628-632 (1973).
- Goodfellow, P. *et al. Eur. J. Immun.* **12**, 627-704 (1982).
- Köhler, G. & Milstein, C. *Nature* **256**, 495-497 (1975).
- Povey, S. *et al. Ann. Hum. Genet.* **43**, 241-248 (1980).
- Solomon, E. *et al. Somatic Cell Genet.* **2**, 125-140 (1976).
- Hope, R. M., Goodfellow, P. N., Solomon, E. & Bodmer, W. F. *Cytogenet. Cell Genet.* **33**, 204-212 (1982).
- Goodfellow, P. N., Banting, G., Trowsdale, J., Chambers, S. & Solomon, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1190-1194 (1982).
- Solomon, E. *et al. Cytogenet. Cell Genet.* (in the press).
- Schweizer, D., Ambros, P. & Andrie, M. *Expl Cell Res.* **111**, 327-332 (1978).

The membrane form of variant surface glycoproteins of *Trypanosoma brucei*

M. Lucia Cardoso de Almeida & Mervyn J. Turner
MRC Biochemical Parasitology Unit, The Molteno Institute,
Downing Street, Cambridge CB2 3EE, UK

African trypanosomes are parasitic protozoa which are enveloped by a surface coat consisting of a matrix of identical glycoprotein molecules¹. Variations in the composition of these variant surface glycoproteins (VSGs) allow the parasite to escape the host's immune system and render effective immunophylaxis improbable²⁻⁵. However, underlying the surface coat, all variant antigen types contain common membrane components⁶, some of which can activate complement by the alternative pathway⁷, leading to lysis of uncoated trypanosomes. Hence, stimulation of VSG release *in vivo* should be a potential form of chemotherapy, and we have therefore investigated the mode of attachment of VSG to the plasma membrane. Biochemical characterization of VSGs from several species has been performed on material purified after release from the cell surface following rupture of the trypanosome⁸⁻¹¹. We demonstrate here that VSGs of *Trypanosoma brucei* when bound to the membrane exist in a form which differs both biochemically and immunochemically from VSGs purified in the conventional manner. After rupture of the cell, membrane-form VSG (mfVSG) is enzymatically transformed into the commonly isolated water-soluble released form (sVSG). In conditions in which this modification does not take place, purified VSGs have amphiphilic properties and behave as

integral membrane proteins by the criterion of charge-shift electrophoresis. The difference between the two forms lies in the C-terminal domain, which is phosphorylated in both forms. This domain in sVSGs contains an immunogenic oligosaccharide known as the cross-reacting determinant (CRD), attached to the C-terminal amino acid¹². Recognition of this determinant by anti-CRD antibodies is impaired in the membrane form.

VSGs can be purified to homogeneity following their release in a water-soluble form from ruptured trypanosomes⁸. It has been commonly assumed that these purified VSGs represent the form bound to the membrane. However, by using SDS-polyacrylamide gel electrophoresis (PAGE), purified sVSG can be distinguished from the VSG present on trypanosomes that have been lysed by boiling in detergent (Fig. 1A,B). Surprisingly, SDS-PAGE revealed that sVSGs have an apparent molecular weight 800-2,000 daltons higher than VSGs in detergent lysate. Furthermore, after electrophoretic transfer of the proteins from the gel to nitrocellulose paper, binding of the anti-CRD antibody could be detected more readily to the purified water-soluble VSGs (Fig. 1B). *T. brucei* variant MITat 1.5 shows low-affinity binding of purified VSG to anti-CRD (Fig. 1A,B; track 1b) consistent with results obtained in quantitative radioimmunoassay (data not shown).

This molecular weight change is associated with the release of soluble VSGs from the trypanosome (Fig. 1C). SDS-PAGE showed that before rupture all the VSG was in the high-mobility form. After one cycle of freeze-thawing, a mixture of the two forms was observed, and following centrifugation, the supernatant contained only VSG in the low-mobility form, which co-migrated with purified VSGs.

Analysis of these events in several variants led us to conclude that the conversion was a dynamic process which could be inhibited by boiling in detergent. Preparation of lysates in neutral detergents at lower temperatures did not inhibit the conversion from low- to high-mobility form (see Fig. 2). Conversion occurred faster at 23 °C than at 4 °C, but even after 45 min at 23 °C, a small amount of the membrane form was still detectable. The rate of conversion at 4 °C was perfectly compatible with the time scale of sVSG preparation by conventional means⁸. This observation provides an explanation for the variation in the amount of antigen remaining in the membrane after rupture of the trypanosome^{3,8,13} and the finding that multiple washes of the insoluble fraction following breakage of the trypanosomes greatly increases the yield of water-soluble VSG³. Figure 2C clearly illustrates the striking difference in affinity of anti-CRD antibody for the two forms. The same difference in affinity has also been observed by quantitative radioimmunoassay and solution immunoprecipitation (data not shown). Figure 2D,E shows that a VSG can be converted from the high-mobility to low-mobility form by incubation with *n*-octyl- β -D-glucopyranoside (nOG) lysates from heterologous clones of trypanosomes if such lysates are prepared at 23 °C, but not at 100 °C. Clearly, the data are compatible with the existence of an enzyme (or enzymes) which can act on all VSGs, and which has as its substrate a feature common to all membrane-form VSGs.

We observed that trypanosomes subjected to hypotonic lysis in the presence of 10 mM zinc chloride retained their surface coat (data not shown) and that in these conditions the molecular weight change was inhibited (Fig. 3A,B). This observation, consistent with the report that zinc ions inhibit VSG release^{14,15}, allowed the solubilization of VSG in the high-mobility form (Fig. 3C,D), and subsequent purification to homogeneity as described in Fig. 4 legend. The detergent-solubilized VSG retained its apparently lower molecular weight throughout the purification process (Fig. 4A).

To characterize this VSG as an integral membrane protein, we used charge-shift electrophoresis¹⁶. Detergent binding to hydrophobic proteins can be detected by the use of ionic detergents; the competitive binding of the latter in the presence of non-ionic detergents alters the net charge on the protein, and

Fig. 1 Purified VSGs and VSG in detergent lysates have different apparent molecular weights on SDS-PAGE; this difference is produced during breakage of trypanosomes. Trypanosome lysates were prepared by boiling 2.5×10^8 cells ml^{-1} in phosphate-buffered saline (PBS) containing 2% SDS. After cooling, 2.5% Triton X-100 in PBS was added to give a final concentration of 2% Triton X-100 and 0.4% SDS²³. Samples were run on 7–20% polyacrylamide gradient gels in the presence of 0.1% SDS, using the buffer system of Laemmli²⁴. **A**, Coomassie blue staining pattern of a gel loaded with lysates (tracks a) and purified sVSGs (tracks b) prepared from *T. brucei brucei* variants MITat 1.5, 1.2, 1.7, 1.6 and ILTat 1.25, 1.26 and 1.4 (tracks 1–7, respectively). **B**, Electrophoretic transfer to nitrocellulose paper of gel shown in **A**, performed using a modification of the method of Towbin *et al.*²⁵. The paper, after blocking with excess protein, was incubated overnight with purified anti-CRD antibody. This was obtained from the serum of a rabbit immunized with ILTat 1.21, then affinity-purified, first on ILTat 1.21 sVSG coupled to Sepharose 4B, and then on MITat 1.6 sVSG-Sepharose 4B. The eluate from this second column cross-reacts with all VSGs tested to date by radioimmunoassay. Antibody bound to the nitrocellulose was detected using ¹²⁵I-labelled-goat anti-rabbit immunoglobulin and the radiolabel was visualized by autoradiography. Identical results were obtained in the presence or absence of Triton X-100; in mixtures of purified VSG and its respective lysate the difference can still be observed (data not shown). **C** Shows that this change in apparent molecular weight of the VSG occurs on rupture of the cells. A suspension of trypanosomes (MITat 1.2 at $5 \times 10^6 \text{ ml}^{-1}$) in PBS was lysed by freezing in liquid nitrogen and thawing at room temperature. One such treatment ruptured all the trypanosomes. After centrifugation at 105,000g at 4°C for 1 h, the supernatant was removed and the pellet washed once with PBS. Samples equivalent to an initial value of 3×10^6 trypanosomes were prepared at each stage by immediately boiling in SDS-gel sample buffer, and analysed by SDS-PAGE. Tracks 1, 10, molecular weight standards; tracks 2, 9, purified MITat 1.2 sVSG; track 3, sample prepared from trypanosomes before freeze-thawing; track 4, after thawing; track 5, supernatant after centrifugation; track 6, pellet after centrifugation; track 7, supernatant derived from washing the pellet in 6; track 8, final pellet. Identical results were obtained for several variants tested in the same conditions. The 67,000 (67K)-molecular weight band in **C** is albumin.

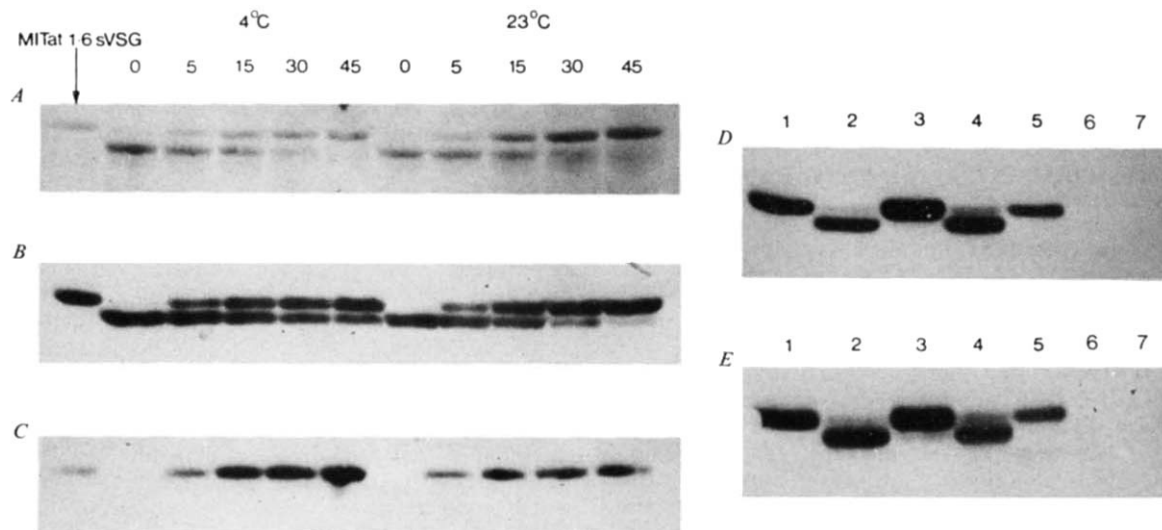
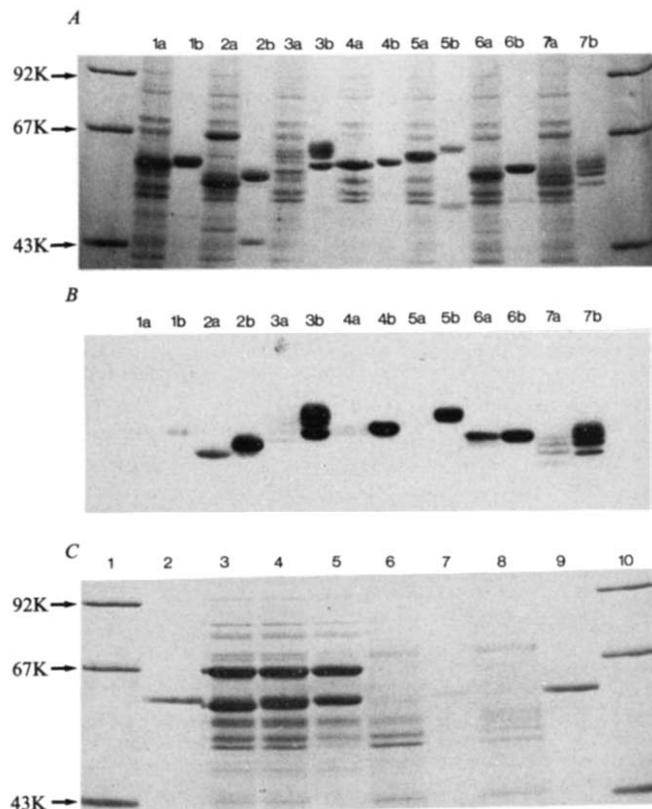


Fig. 2 Fast-migrating VSG is converted to the slow-migrating form within 45 min and this conversion can be mediated by heterologous lysates. MITat 1.6 trypanosomes ($2 \times 10^6 \text{ ml}^{-1}$) were lysed in PBS containing 1% nOG at 4°C and 23°C. At $t = 0, 5, 15, 30$ and 45 min of incubation, aliquots were removed and immediately boiled in SDS-gel sample buffer. Samples equivalent to 4×10^6 trypanosomes were fractionated by SDS-PAGE in the presence of 4 M urea, which was found to greatly enhance the resolution between the two forms of VSG. Replicas of this gel were transferred to nitrocellulose paper to allow immunochemical identification. **A** Shows the Coomassie blue staining; **B**, MITat 1.6 bands identified using a monoclonal anti-MITat 1.6 antibody (M10A) and ¹²⁵I-rabbit anti-mouse immunoglobulin. **C**, A similar nitrocellulose transfer was incubated with anti-CRD antibody and bound antibody detected using ¹²⁵I-goat anti-rabbit immunoglobulin. Purified MITat 1.6 sVSG was applied to gels as a standard. To assess whether this conversion of the VSG molecule could be effected by enzymes present in *T. b. brucei* of different variant antigen types, lysates of MITat 1.6 trypanosomes were prepared in 1% nOG/PBS at 100°C or 23°C. The VSG present in such lysates was then detected in SDS-gels run in the presence of 4 M urea, after electrophoretic transfer to nitrocellulose, using monoclonal anti-MITat 1.6 VSG antibody and ¹²⁵I-rabbit anti-mouse immunoglobulin. Heterologous lysates of ILTat 1.25 and MITat 1.7 trypanosomes were prepared at the same time in exactly the same conditions. MITat 1.6 lysate prepared at 100°C was incubated at 23°C for 30 min with an equal volume of the homologous and heterologous lysate, prepared at either 100°C or 23°C. **D**, The heterologous lysate used was MITat 1.7; **E**, ILTat 1.25. Track 1, purified MITat 1.6 sVSG; track 2, MITat 1.6 lysate prepared at 100°C; track 3, MITat 1.6 lysate prepared at 23°C; tracks 4 and 5, 100°C MITat 1.6 lysate incubated with heterologous lysate prepared at 100°C and 23°C, respectively; tracks 6 and 7, heterologous lysates alone, prepared at 100°C and 23°C, respectively. All the samples were incubated at 23°C for 30 min, then boiled in SDS-gel sample buffer and subsequently fractionated on SDS-PAGE.

Fig. 3 Change of apparent molecular weight is inhibited by zinc ions, allowing solubilization of VSG in the fast-migrating form. Trypanosomes (MITat 1.6 at 2×10^6 ml $^{-1}$) were submitted to hypotonic lysis in water (A) or in the presence of 10 mM ZnCl₂ (B), followed by one cycle of freezing and thawing as described in Fig. 1 legend. Soluble and insoluble VSGs were separated by centrifugation at 105,000g for 1 h at 4°C, and samples were immediately prepared for SDS-PAGE by boiling in sample buffer. Tracks 1, 6, molecular weight standards including MITat 1.6 sVSG. Track 2, trypanosomes following hypotonic lysis; track 3, following freeze-thawing; track 4, supernatant after high-speed centrifugation; track 5, pellet. C and D show the results of the same procedures done in 50 mM HEPES-NaOH, pH 6.9, 1% nOG, in the presence (D) or absence (C) of 2 mM ZnCl₂. The gel was stained with Coomassie brilliant blue.

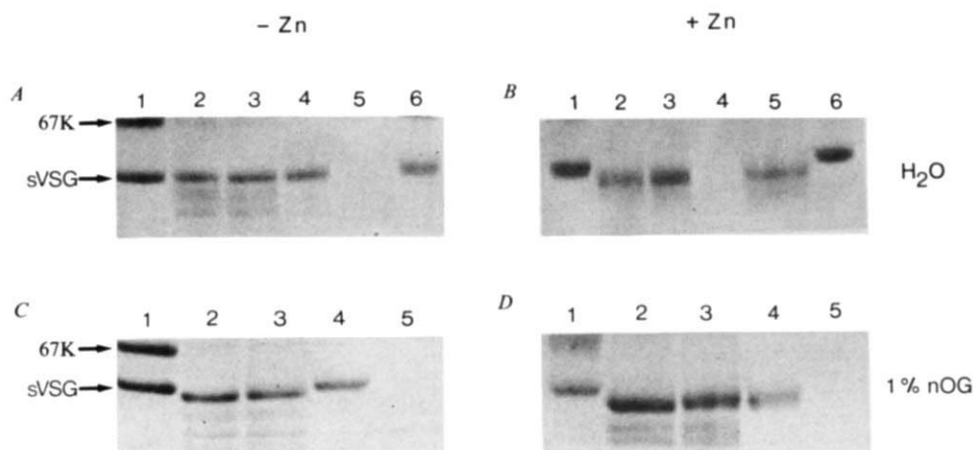
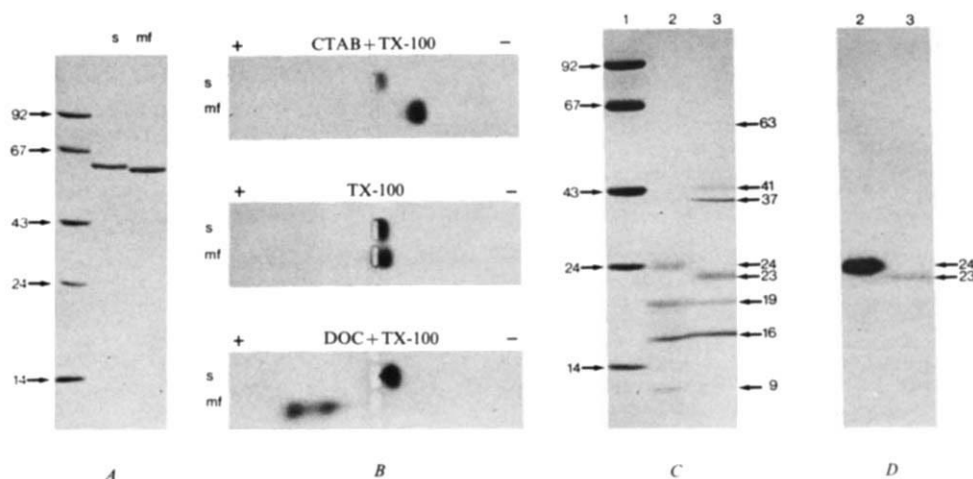


Fig. 4 MITat 1.6 mfVSG was purified to homogeneity and compared with sVSG by charge-shift electrophoresis. The apparent molecular weight difference lies in the C-terminal CNBr peptide containing the CRD. Water-soluble MITat 1.6 sVSG was prepared according to the method of Cross⁸. Detergent-soluble MITat 1.6 mfVSG was prepared by resuspending a pellet containing 1.3×10^8 trypanosomes in 1 ml of 50 mM HEPES-NaOH pH 6.9, 1% nOG, 0.1 mM phenylmethylsulphonyl fluoride, followed immediately by centrifugation at 105,000g for 1 h at 4°C. The supernatant was applied to a 1-ml column of affinity-purified rabbit α MITat 1.6 VSG coupled to Sepharose 4B. After washing with solubilization buffer, bound protein was eluted with 1 M propionic acid containing 1% nOG. SDS-PAGE showed that the eluate contained the VSG in homogeneous fast-migrating form (A). Samples (40 μ g) for charge-shift electrophoresis were freeze-dried and redissolved in 30 μ l of fivefold concentrated running buffer (see below) and were incubated for 2 h at room temperature. Immediately before application to the gel, 2 μ l of 0.05% bromophenol blue/50% glycerol was added, and samples were electrophoresed overnight at 100 V, 130 mA in a 1% horizontal agarose gel, prepared in running buffer. The gel was cooled, and buffer recirculated continuously. Running buffer was 0.05 M glycine-NaOH pH 9.0, 0.1 M NaCl, 0.5% Triton X-100 (TX-100) used either alone or supplemented with 0.05% cetyltrimethylammonium bromide (CTAB) or 0.25% sodium deoxycholate (DOC) as described elsewhere¹⁶. B, Coomassie blue-stained electrophoretogram of these samples; the anode is at the left and the cathode at the right of each gel. Taking advantage of the known CNBr fragmentation pattern of MITat 1.6 sVSG (ref. 23), the location of the difference responsible for the altered mobility on electrophoresis was determined. Both forms of purified VSG were digested with CNBr as described elsewhere²⁶ and the products were analysed by SDS-PAGE. C, Coomassie blue-stained gel: track 1, molecular weight standards; track 2, sVSG CNBr peptides; track 3, mfVSG CNBr peptides. A replica of this gel on nitrocellulose paper was incubated with anti-CRD antibody. The autoradiogram of this replica is shown in D; tracks 2 and 3 as in C.



therefore its electrophoretic mobility. Comparing both forms of MITat 1.6 VSG (Fig. 4A), only the mobility of the detergent-soluble form was affected by the presence of ionic detergent (Fig. 4B). This behaviour, which is characteristic of integral membrane proteins, is compatible with the role of VSG in forming a functional surface coat. As shown in Figs 1 and 3, all VSG molecules can be solubilized in the high-mobility form, and all of them behave as integral membrane proteins, indicating that mfVSG is not a subset of the total VSG population, such as secreted and bound immunoglobulins¹⁷, but rather that sVSG is a derivative of the amphiphilic form bound to the membrane. The reason for the heterogeneous nature of detergent-binding mfVSG in sodium deoxycholate (DOC) is unknown.

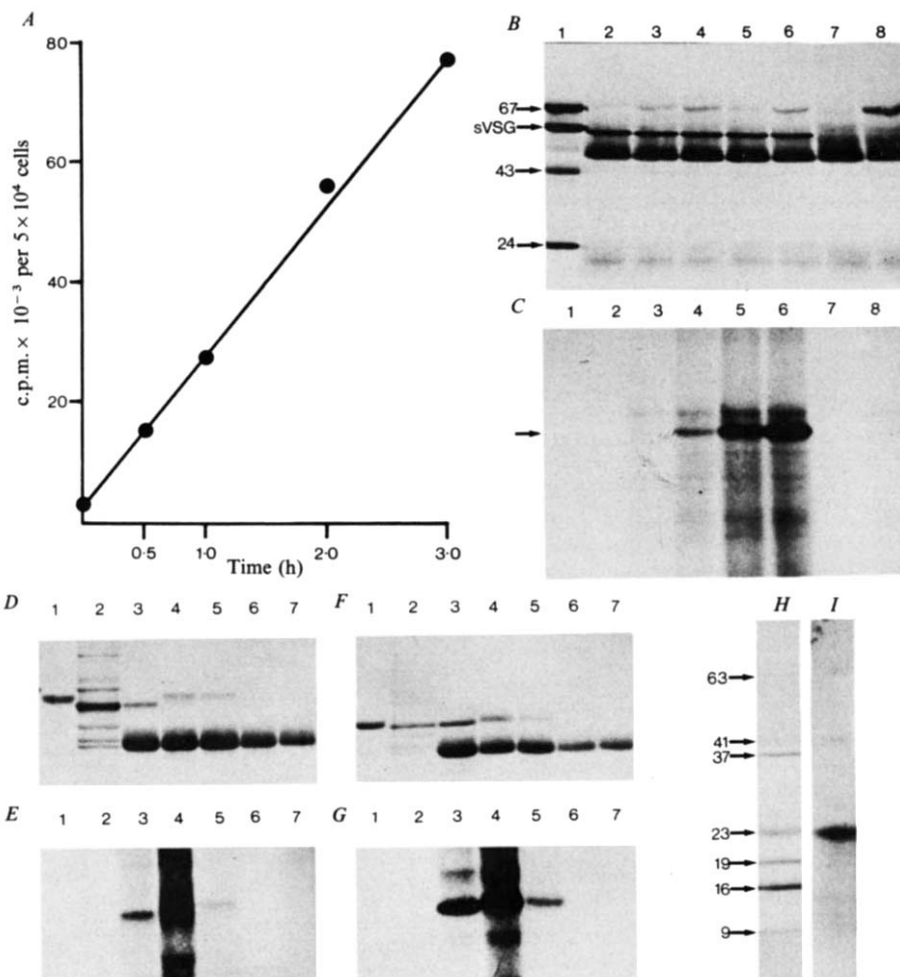
The molecular weight difference between MITat 1.6 mfVSG and sVSG was shown to reside in the C-terminal CNBr peptide containing the CRD (Fig. 4C,D). This suggests the presence of a covalent modification, which affects the structure and/or conformation of the CRD, and whose presence is presumably essential for membrane binding.

We considered the possibility that phosphorylation/dephosphorylation controls the release of VSGs from trypanosomes.

Figure 5A shows that ³²P-phosphate was readily incorporated into mfVSG and the radiolabel was present exclusively in CNBr peptides containing the C-terminus (Fig. 5H,I). However, after one cycle of freeze-thawing, sVSG was also phosphorylated. In the absence of further information about the phosphate content of the two forms, we cannot exclude the possibility that they are differentially phosphorylated. Work is in progress to resolve this point.

The nature of the difference between sVSG and mfVSG is at present unknown. One immediate possibility is the presence in mfVSG of the hydrophobic C-terminal extension whose existence has been inferred from cDNA sequence analysis¹⁸⁻²⁰, but which has never been observed on isolated VSGs. This can only be finally resolved by C-terminal sequence analysis. At present, we are attempting to purify sufficient mfVSG to do this. However, two lines of evidence suggest that the hydrophobic tail is absent. First, pulse-labelling experiments suggest that the hydrophobic tail is processed before VSG is expressed at the plasma membrane²¹. Second, recent work has shown that the CRD is attached to sVSG through an ethanolamine linkage to the α -carboxyl group of the C-terminal amino acid²². As we believe that the CRD is present on mfVSG, albeit in a modified

Fig. 5 Both forms of VSGs are phosphorylated. Trypanosomes were cultured at 37°C at a concentration of $5 \times 10^6 \text{ ml}^{-1}$ in phosphate-free Dulbecco's modified Eagle's medium, supplemented with 10% heat-inactivated dialysed rabbit serum, containing 0.5–1.0 mCi ml^{-1} carrier-free ^{32}P -orthophosphate (Amersham). **A**, Incorporation of ^{32}P into TCA-precipitable material as a function of time. Samples of trypanosomes withdrawn at $t = 0, 0.5, 1, 2$ and 3 h were washed and SDS-Triton X-100 lysates, as described in Fig. 1 legend, were prepared. VSG in such lysates was immunoprecipitated using 50 μg of affinity-purified homologous antibody or normal rabbit serum, and heat-killed formalinized *Staphylococcus aureus* Cowan strain A²⁷. Immune complexes were released by boiling in SDS-gel sample buffer and were analysed by SDS-PAGE followed by autoradiography. **B**, Coomassie blue-stained gel obtained in such an experiment using MITat 1.6 trypanosomes. **C**, The autoradiogram of this gel. Arrow in **C** indicates the migration position of mfVSG. Track 1, molecular weight standards including purified MITat 1.6 sVSG; tracks 2–6, immune complexes isolated after labelling for 0, 0.5, 1, 2 and 3 h respectively; tracks 7 and 8, normal rabbit serum controls after 1 and 3 h of labelling, respectively. In separate experiments, trypanosomes radiolabelled for 3 h were washed, resuspended in PBS, and submitted to one cycle of freeze-thawing. Released and bound VSGs were separated by centrifugation at 200,000g for 30 min at 4°C. Samples from each stage were immunoprecipitated from 1% nOG lysates prepared at 100°C, with the homologous antiserum as described above. This experiment was performed using ILTat 1.25 (**D,E**) and MITat 1.6 (**F,G**) trypanosomes radiolabelled for 3 h; the Coomassie blue-stained gels are shown (**D,F**) and the corresponding autoradiograms (**E,G**). Track 1, purified sVSG; track 2, whole trypanosome lysate, prepared before labelling; tracks 3 and 7, immune complexes precipitated from radiolabelled lysates after incubation with homologous antiserum or normal rabbit serum, respectively; tracks 4–6, immune complexes isolated after incubation with homologous antiserum following freeze-thawing from the supernatant and from the pellet, respectively. **H,I**, A lysate from 5×10^6 radiolabelled MITat 1.6 trypanosomes was prepared by boiling in 1% nOG in PBS and was centrifuged at 200,000g for 30 min at 4°C. The supernatant was absorbed and eluted from a column of affinity-purified rabbit anti-MITat 1.6 sVSG antibody coupled to Sepharose 4B, as described in Fig. 4 legend. The eluate was freeze-dried, digested with CNBr and the products analysed by SDS-PAGE. **H**, Coomassie blue-stained gel; **I**, autoradiogram of the same gel.



form, mfVSG cannot also contain the hydrophobic tail. Other possibilities, such as the presence of covalently bound lipid or phospholipid, are now being investigated. We believe that the difference between mfVSG and sVSG reflects the presence of a modification in the mfVSG which alters the SDS-binding characteristics of the protein, and thus changes its migration pattern on polyacrylamide gels.

In all species of salivarian trypanosomes examined to date, VSGs are readily released (for review see ref. 3), suggesting that the same or similar release mechanisms operate across species barriers. Whether this mechanism has a specific role in the life cycle of the trypanosome remains a matter for conjecture. It could be involved in the turnover of VSG molecules, or in the transformation of coated bloodstream trypomastigotes to coatless procyclic trypomastigotes in the tsetse fly midgut. It could also be a nonspecific degradative event triggered by breakage of the cells. For these reasons, characterization of the

enzyme and information concerning its distribution within the life cycle of the trypanosome are urgently required. We believe that investigation of the events modulating this mechanism may lead to the development of new approaches to the control of trypanosomiasis.

We thank Drs R. W. LePage and D. W. Taylor for assistance in establishing optimal conditions for ^{32}P -labelling; Dr E. N. Miller and L. Allan for providing monoclonal antibodies; Josie McConnell for helpful discussions; and Dr A. A. Holder for communication of results before publication. We also thank Esther Choolun, Judith Creighton, Roger Williams and David Law for technical assistance. M.L.C. de A. is supported as a postgraduate research student by FAPESP and CNPq from Brazil; this work will form part of her PhD dissertation. This investigation received financial support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases.

Received 8 October 1982; accepted 19 January 1983.

- Vickerman, K. J. *Cell Sci.* **5**, 163–193 (1969).
- Vickerman, K. *Nature* **273**, 613–617 (1978).
- Turner, M. J. *Adv. Parasit.* **21**, 69–153 (1982).
- Englund, P. T., Hajduk, S. L. & Marini, J. C. A. *Rev. Biochem.* **51**, 695–726 (1982).
- Borst, P. & Cross, G. A. M. *Cell* **29**, 291–303 (1982).
- Seed, J. R. *Parasitology* **54**, 593–596 (1964).
- Tetley, L., Vickerman, K. & Moloo, S. K. *Trans. R. Soc. Trop. Med. Hyg.* **75**, 409–414 (1981).
- Cross, G. A. M. *Parasitology* **71**, 393–417 (1975).
- Olenick, J. G., Travis, R. W. & Garson, S. *Molec. Biochem. Parasit.* **3**, 227–238 (1981).
- Baltz, T. et al. *FEBS Lett.* **82**, 93–96 (1977).
- Reinwald, E., Rautenberg, P. & Risse, H. J. *Biochim. biophys. Acta* **668**, 119–131 (1981).
- Holder, A. A. & Cross, G. A. M. *Molec. Biochem. Parasit.* **2**, 135–150 (1981).
- Rovis, L., Barbet, A. F. & Williams, R. O. *Nature* **271**, 638–656 (1978).
- Bowles, D. J. & Voorheis, H. P. *FEBS Lett.* **139**, 17–21 (1982).

- Voorheis, H. P., Bowles, D. J. & Smith, G. A. *J. biol. Chem.* **257**, 2300–2304 (1982).
- Helenius, A. & Simons, K. *Proc. natn. Acad. Sci. U.S.A.* **74**, 524–532 (1977).
- Parkhouse, R. M. E., Lifter, J. & Choi, Y. S. *Nature* **284**, 280–281 (1980).
- Boothroyd, J. C., Cross, G. A. M., Hoijmakers, J. H. J. & Borst, P. *Nature* **288**, 624–626 (1980).
- Matthysens, G., Michiels, F., Hamers, R., Pays, E. & Steinert, M. *Nature* **293**, 230–233 (1981).
- Rice-Ficht, A. C., Chen, K. K. & Donelson, J. E. *Nature* **294**, 53–57 (1981).
- McConnell, J., Turner, M. J. & Rovis, L. *Molec. Biochem. Parasit.* (in the press).
- Holder, A. A. *Biochem. J.* **209**, 261–262 (1983).
- McConnell, J., Gurnett, A. M., Cordingley, J. S., Walker, J. E. & Turner, M. J. *Molec. Biochem. Parasit.* **4**, 225–242 (1981).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Towbin, H., Staehlin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).
- Cross, E. *Meth. Enzym.* **11**, 238–255 (1967).
- Kessler, S. W. *J. Immun.* **115**, 1617–1624 (1975).

Steady-state magma discharge at Etna 1971–81*

ESTIMATIONS by Wadge and Guest¹ of the volumes of lava emitted by each individual eruption of Mount Etna between 1971 and 1981, should be viewed with caution^{2,3}. Despite the impressive number of references (see Table 1 in ref. 1), very few of them contain accurate determinations of the volumes of lava erupted, that is: measurements of the surface covered by each lava flow and a sufficient number of determinations of the thickness. Regarding this latter point, it is symptomatic that all but one (J. B. Murray) of the authors to whom reference is made fail to mention any thickness determination, so that the volumes quoted are in most cases highly speculative: they are not measurements in a scientific sense⁴.

Let us now examine specific points:

- 5 April–12 June 1971: since the error bar of the volume estimate of this flank eruption is unknown, there is no means of inferring that the accuracy is 'probably' better than 50%. For another flank eruption (1950–51), calculations by J. B. Murray (personal communication) from the 1932 and 1969 IGM maps lead to a value of $116 \times 10^6 \text{ m}^3$ ($\pm 10\%$), instead of $55 \times 10^6 \text{ m}^3$ as 'estimated' by Wadge².

- June–September 1971/October 1973: no lava effusion occurred during this period. It is difficult to imagine the meaning of the $17 \times 10^6 \text{ m}^3$ listed. If it represents the filling of the Central Crater Chasm, it must be pointed out that the dimensions of its interior were not known³.

- Eruptions from the NE Crater and the northern flank, 1974–76: except for the small area ($800 \times 500 \text{ m}$) around the NE Cone^{5,6} and for a small part of the lava field (very approximate sketch in ref. 7), there is no map of the erupted products. How, therefore, can their volume be calculated?

- Eruptions from the NE Crater, 1977–78: we have a succession of 17 eruptions, most of which occurred at intervals of a few days in a zone of the volcano which is almost inaccessible during the winter. Not only is there no available map, but the exact location of many flows is not even known. Similar observations can be made for the 1980–81 eruptions of the same NE Crater: the authors to whom reference is made¹ indicate only the approximate length of the lava flows.

- Flank eruptions, 1978–79: none of these flows has been mapped so far. The

sketch shown by Mackey and Scott⁸ only represents a small part of the flow corresponding to the less important of the four eruptions.

- March 1981 flank eruption: this is the only case where a detailed map has been made⁹ and a volume actually measured. It represents less than 10% of the total volume taken into account by Wadge and Guest¹ during the period 1971–81.

Therefore, I consider the conclusions drawn by Wadge and Guest to be unfounded, unless they can display a precise topographic map for each of the 35 eruptions listed in their Table 1. Although a steady-state magma discharge appears obvious to many observers for some periods of Mount Etna's 'persistent activity', it cannot be quantified with such speculative estimates. The observations by Italian, British and French researchers

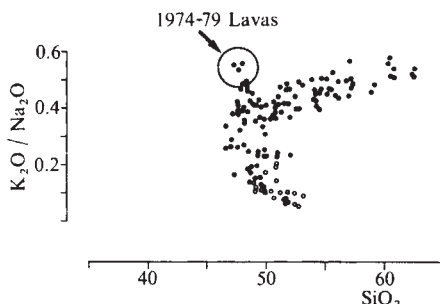


Fig. 1 Composition of the 1974–79 lavas compared with that of all the products of Mt Etna ($\text{K}_2\text{O}/\text{Na}_2\text{O}$ versus SiO_2). ○, Pre-eruptive tholeiitic basalts; ●, alkaline lavas from the strato-volcano of Etna itself, including historic flows (from ref. 4).

only permit the definition of an order of magnitude and we must await extensive topographic work to know with a sufficient degree of accuracy the volume of lava erupted in the decade 1971–81.

Finally, it is incorrect to emphasize the chemical homogeneity of recent lavas of Etna. Since 1974, a significant change in the composition of the products has occurred, with an evolution towards a more potassic character (see Fig. 1)⁴. Added to the exceptional frequency of the eruptions in recent times, this observation seems to indicate a new magma supply from a deep source and does not allow us to guess the future behaviour of Mount Etna.

J. C. TANGUY

Université de Paris 6 (PIRPSEV),
4, avenue de Neptune,
94107 St Maur des Fossés Cedex, France

1. Wadge, G. & Guest, J. E. *Nature* **294**, 548–550 (1981).
2. Wadge, G. *J. Volcan. geotherm. Res.* **2**, 361–384 (1977).
3. Tanguy, J. C. *J. Volcan. geotherm. Res.* **6**, 179–188 (1979).
4. Tanguy, J. C. thesis, Univ. Paris 6, PIRPSEV (1980).

5. Murray, J. B., U.K. *Research on Mt Etna 1975–76*, 9–12 (Royal Society, London, 1977).
6. Murray, J. B. U.K. *Research on Mt Etna 1977–79*, 33–42 (Royal Society, London, 1980).
7. Pinkerton, H. & Sparks, R. S. J. *J. Volcan. geotherm. Res.* **1**, 167–182 (1976).
8. Mackey, M. & Scott, S. C. U.K. *Research on Mt Etna 1977–79*, 43–44, (Royal Society, London, 1980).
9. VEST investigations PIRPSEV Bull. No. 57 (INAG, Paris, 1982).

WADGE AND GUEST REPLY—Tanguy argues that our interpretation of magma discharge at Etna during 1971–81 is invalid because our estimates of lava volumes are too imprecise. His principal argument is that detailed maps and measured lava thicknesses are not available in most of the references quoted by us. We would point out, however, that our volume estimates are not based on information from these references alone, and that one of us has visited the volcano on average twice a year during the past decade and wherever possible has mapped the areal extent of new lavas and made thickness measurements. These data were used in our calculation of lava volumes and indeed are in part the basis of our study.

Thus, although Tanguy states that the 1981 eruption is the only one where detailed mapping and thickness measurements were made (by one of us and other UK workers), similar measurements were made on flows from other major eruptions. Where we were not able to obtain such data, especially for short-lived NE Crater eruptions, we have relied on eye-witness reports of others, for example R. Romano (personal communication) for the July 1977 eruption and C. R. J. Kilburn (personal communication) for the September 1980 eruption. The eruptions mentioned for June and September 1971 and October 1973 were indeed eruptions on the floor of the Chasm and the Bocca Nuova. Both these eruptions occurred when the internal dimensions of these two pits in the central crater were known from visual observations and range-finder measurements. Our original letter was an inappropriate place to justify each estimated volume.

Tanguy appears to accept our model of a steady state magma discharge but rejects our volumetric data as 'speculative estimates'. However, he does not present any new data that demonstrate our estimates to be less precise than stated in our letter, and indeed the only quantitative data given by him are for an eruption that occurred in 1950–51 which is outside the period of our study.

With regard to Tanguy's comments on lava chemistry, each of the erupted lavas was hawaiitic in common with most historical lavas; but as we pointed out, there were limited variations in composition. Some of these occurred within single eruptions including those of 1971¹ and 1981 (S. Scott, personal communication).

*This Matters Arising exchange replaces that published previously (13 January, p. 175) in which some extraneous text was mistakenly substituted in the "reply". We apologize to Drs Wadge and Guest for this error.

when chemically zoned magma-filled fissures may have developed during short repose periods before eruption. Other small variations in chemistry between one set of eruptions and another have been observed during the last decade and may identify individual batches of magma with different chemical signatures (S. Scott, personal communication). The careful monitoring of Etna's output is clearly important in interpreting these chemical observations in terms of the volcano's internal plumbing system.

In our letter we made the empirical observation that the cumulative volume curve defined an envelope of steady-state behaviour. We suggested that because the curve would leave the envelope during May 1982 $\pm$ 3 months an eruption before this time could be expected. In fact a glow was observed on the floor of the Bocca Nuova in February 1982 and incandescent scoria rose above the crater rim on 20–22 May². Continued strombolian activity was particularly marked on 8 August³ and has continued to the time of writing.

G. WADGE
J. E. GUEST

University of London Observatory,
Mill Hill Park, London NW7 2QS, UK

1. Romano, R. & Sturiale, C. *Phil. Trans. B. Soc. A274*, 37–43 (1973).

2. *SEAN Bull.* 7/5, 12 (Smithsonian Institution, 1982).

3. *SEAN Bull.* 7/7, 8 (Smithsonian Institution, 1982).

Termites and methane

RASMUSSEN AND KHALIL¹ measured laboratory methane emission rates of one species of termite, extrapolated this to a global methane production estimate and concluded that the estimate made by Zimmerman *et al.*² was inflated. The procedures used by Rasmussen and Khalil have several similarities to our research, but with significant differences. It is critical to explain these differences so that our respective estimates of the potential emissions from termites can be evaluated properly.

Rasmussen and Khalil used termite emission rates as the basis for their global extrapolation of methane production. Their termites were sealed in 0.95 l chambers at "room temperature" for up to 4 h and the accumulation of methane was measured. In these conditions humidity would change from ambient levels to 100%, CO₂ levels would build up rapidly and oxygen levels would decrease. Termite metabolism and emission rates can change over an order of magnitude with changes in temperature, humidity, type of substrate, amount of substrate, number of individuals or CO₂ concentration^{3,4}.

Our measurements were made on termites kept in flow-through chambers, and

temperature and humidity were kept constant. Even in these constant conditions we found substantial differences in emission rates among termite colonies and species.

We obtained 15 specimens of *Zootermopsis angusticollis* (Hagen) from Rasmussen and Khalil and repeated their experiment using wood as a substrate and a 3-h enclosure time. The emission rate at 20 °C was about 2.4 μ g per termite per day (2.7 times higher than they reported). The experiment was repeated at 25 °C after the termites had been conditioned at the new temperature for 24 h. The new emission rate was 5.7 μ g per termite per day (more than 6 times higher than they reported). When the termites were put in our flow-through system (3 ml min⁻¹) their emission rates were 3.5 and 6.0 μ g per termite per day at 20 and 25 °C respectively. The emission rate of a similar species [*Zootermopsis nevadensis* (Hagen)] was 7.8 μ g per termite per day at 20 °C in our flow-through system.

Some of the difference between the emission rates reported by Rasmussen and Khalil and those that we measured may have been due to individual colony differences. However, we also found that high CO₂ levels inhibit methane production. The inhibition persists for several hours after the culture has been ventilated, and may have occurred in Rasmussen and Khalil's experiment because they had more termites in their static cultures and their enclosure time was "up to 4 h".

Because of this extreme variability, global extrapolations based solely on emission rates are unreliable. In our paper² we used the ratio of the total gas evolved over a 3 or 4 week period to the food consumed by the termites as the basis for our extrapolation. These emission efficiencies have been measured over wide ranges of termite numbers and CO₂ concentrations and are much more uniform among various termite species than emission rates. For example, the efficiency of methane production for *Reticulitermes tibialis* (Banks) was similar at 30 °C and at 20 °C although emission rates varied by more than a factor of 10. In fact the production efficiency of the *Zootermopsis* species obtained from Rasmussen and Khalil and measured over a 41-day period was the same as that reported in our paper for *Reticulitermes* and *Gnathamitermes*². We hypothesize that although termite activity in the laboratory is no doubt substantially different than for intact colonies in the field, the basic stoichiometry of the digestive process is not affected as greatly as emission rates.

Rasmussen and Khalil's equation (3) used to describe our method to estimate global production of methane by termites, is misleading. We did not use estimates of the fraction of the net primary productivity (NPP) consumed by termites in each ecosystem directly in our calculations.

Instead we multiplied termite density estimates by a consumption figure which was the average of many consumption estimates for many termite species reported in the literature. The biomass consumed was then multiplied by the efficiencies (emission yields per gramme carbon ingested) which we determined in laboratory studies.

We have found no literature support for the statement in Rasmussen and Khalil's paper (referring to termite density estimates) that "This number may be high since densities are generally measured where there are a lot of termites, and therefore may not reflect the world average for similar environments". Contrary to this statement, in many cases termite density estimates have been by-products of research designed to identify soil invertebrates. The authors usually conclude that the densities are probably minimum estimates since relatively shallow soil depths are sampled and termites have been known to penetrate at least 70 m into the soil^{5,6}. We have also made independent energy balance calculations to see if the oxygen required to break down the amount of material estimated to be consumed by the number of termites in our global budget would yield a reasonable metabolic size for termites. The agreement was good². In addition we have estimated termite densities for selected tropical regions where mound builders are prevalent, by multiplying termite mounds per unit area (determined from areal photographs), area-specific literature estimates of the percentage of active nests and estimates of termites per mound. Again the agreement indicated that our estimates were not inflated.

We have been conducting laboratory and fieldwork on emissions from termites for 3 years. Our research has continued since submission of our paper to *Science* in August of 1981. We have now made thousands of measurements on six different termite species in our laboratory and we have sampled emissions from seven different species in the field. The results have been consistent with our published work. We are convinced that although measuring methane accumulation in a closed jar is simple, the data are unreliable for global extrapolations. Our research has shown that methane emission rates can be easily manipulated in the laboratory. Digestive processes, however, seem to be quite reproducible over a wide range of species and experimental conditions. We therefore believe that digestive efficiencies are a sounder foundation on which to base global estimates.

PATRICK R. ZIMMERMAN
JAMES. P. GREENBERG

National Center for
Atmospheric Research,
Boulder, Colorado 80307, USA

1. Rasmussen, R. A. & Khalil, M. A. K. *Nature* **300**, 700 (1983).
2. Zimmerman, P. R., Greenberg, J. P., Wandiga, S. O. & Crutzen, P. J. *Science* **218**, 563–565 (1982).
3. Becker, G. in *The Biology of Termites*, Vol. 1 (eds Krishna, K. & Weesner, F. M.) 351–385 (Academic, New York, 1969).
4. Peakins, G. J. & Josens, G. in *The Biology of Termites* Vol. 1 (eds Krishna, K. & Weesner, F. M.) 111–163 (Academic, New York, 1969).
5. Wood, T. G. & Sands, W. A. in *The Biology of Termites* Vol. 1 (eds Krishna, K. & Weesner, F. M.) 242–292 (Academic, New York, 1969).
6. Baroni-Urbani, C., Josens, G., Peakins, G. J. in *The Biology of Termites* Vol. 1 (eds Krishna, K. & Weesner, F. M.) 5–44 (Academic, New York, 1969).

RASMUSSEN AND KHALIL REPLY—Zimmerman and his colleagues have claimed that termites release 150 Tg of methane into the Earth's atmosphere every year¹. We show in our paper² that this estimate is likely to be too high by at least a factor of 3. The source of our disagreement is the method for extrapolating limited laboratory data to the global environment.

Zimmerman and Greenberg are concerned about high CO₂ levels in our jar studies², however, their use of an open flow-through system may add more uncertainties to measurements of emission rates (or consumption ratios) than a closed jar system. Termite mounds and galleries in the natural environment contain high levels of CO₂ and humidity depending on the types of termites involved and their habitats. Indeed, species such as *Nasutitermes exitiosus* (Hill) and *Coptotermes acinaciformis* (Froggatt) build thick-walled mounds to preserve water vapour and thus CO₂ and CH₄ as well. Peakins and Josens³ show that CO₂ concentration inside the nests of many termite species ranges from 0.4% to 4 or 5% and in certain conditions up to 15%. By comparison atmospheric CO₂ concentration is ~0.03%, and CO₂ measurements in our jar studies were 0.5–1.5%. Thus, a flow-through system may not be representative of the natural environment.

Zimmerman and Greenberg are critical of our equation (3) which was constructed to represent Table 2 of their paper¹. ϵ_i , m_{bi} and A_i are defined by the columns in Table 2 so that their product $\epsilon_i m_{bi} A_i$ is the "total termite consumption"¹¹ = T_i in the i th ecological region. When multiplied by δ , the ratio of CH₄ produced to biomass consumed, one obtains δT_i or the "annual CH₄ production (10¹² g)" in the i th ecological region (Table 2 of ref. 1). Zimmerman and Greenberg state that they calculated T_i by the product $d_i C_i A_i$, where d_i = termite density in the i th ecological region in "Termites per square metre"¹¹, C_i = average biomass consumption (g), and A_i is the area of the region as before (m²). The two formulae are equivalent means for arriving at T_i , except that the average consumption C_i is not given in Table 2 of ref. 1, whereas all the variables of our formula are in the table. We wrote equation (3) to show that T_i is

a product of several variables, each subject to errors which propagate to produce larger errors in estimating T_i . We also wanted to point out that δ , which is the ratio of CH₄ produced to grammes of carbon (biomass) consumed, is assumed to be the same for all ecological regions, all types of diets, and all species of termites. Moreover, it is assumed that δ measured in a few laboratory studies can be safely extrapolated to the varied global environment. Equation (3) may be rewritten as $p_G = \sum_{i=1}^N T_i \delta_i$, which shows that the global production of CH₄ by termites (p_G) is the sum of the methane production in each ecological region (p_i). p_i is the production of CH₄ in the i th ecological region expressed as the product of "total termite consumption"¹¹ (of biomass), T_i and δ_i , the emission yield per gramme carbon ingested which varies from one ecological region to the next. The errors in T_i and δ_i over some 11 ecological regions produce large uncertainties in the calculated p_G . These points regarding error propagation and uncertainty analysis are not affected by the means chosen to calculate T_i (as $\epsilon_i m_{bi} A_i$ or $d_i C_i A_i$). Due to lack of data the true magnitude of the uncertainty in the calculation of p_G cannot be gauged at present, but it is likely to be substantially more than "a factor of 2", which is stated but neither derived nor adequately justified in ref. 1.

This brings us to the crux of the disagreement between us and Zimmerman *et al.* We use measured emission rates ($\bar{p} = \mu\text{g CH}_4$ per yr per termite) to estimate global production as $p_G = \bar{p} N_\infty$ where N_∞ is the total number of termites in the world². Zimmerman *et al.*¹ use measured ratios (δ) of CH₄ emission per gramme carbon ingested and estimate global production as $p_G = \delta T$, where T is the world "total termite consumption" of biomass. With this synopsis of our work and Zimmerman *et al.*'s paper¹, it is clear that claims of one method being superior to the other are difficult to justify since each has its advantages and disadvantages. Is the extrapolation of laboratory measurements of average CH₄ production by termites to global scales any better or worse than extrapolation of laboratory measurements of average CH₄ emission ratio δ to global scales? Is the world population of termites (N_∞) known any more or less accurately than the total biomass consumed (T) by the world's termites? Even if the laboratory measurement of one of the two variables, $\bar{p}$ or δ , is more accurate than the other, do not the enormous uncertainties in N_∞ and T offset any such advantage and still make the two types of estimates equally unreliable?

The claim by Zimmerman and Greenberg that δ is "... much more uniform among various termite species than emission rates" is not only unsupported by scientific studies but may in fact be untrue. For instance, *Coptotermes formosanus* (Shiraki) are a widespread sub-

tropical species which consume enormous amounts of wood (0.98 g (wood) per kg termite per h)⁴, yet produce practically no methane, as shown by Breznak⁵ and in our own experiments ($\bar{p} \sim 0.03 \mu\text{g}$ per termite per day). Wood⁶ has shown that laboratory colonies of *Mastotermes darwiniensis* (Froggatt), *Zootermopsis angusticollis* (Hagen), *Coptotermes acinaciformis* (Froggatt), *Coptotermes lacteus* (Froggatt) and *Nasutitermes exitiosus* (Hill) consume 0.48, 0.42, 0.74, 0.51 and 0.48 g (wood) per kg (termite) per h, respectively. These same termites produce 15, 0.5–0.9, 12, 0.7 and 2.3 mg CH₄ per kg (termite) per h respectively (P. J. Fraser and R. A. R. unpublished data). This amounts to a ratio of δ' of 31, 1–5, 21, 1 and 5 mg CH₄ per kg wood consumed, respectively.

Therefore, Zimmerman *et al.*'s extrapolation of few data for δ to estimate global CH₄ production by termites may be no more reliable than extrapolation of emission rates. In our opinion, the measurement of δ in the natural environment is also more complex and unreliable than an analogous measurement of $\bar{p}$. Field measurement of δ requires an estimate of biomass consumption and termite density, whereas field measurement of $\bar{p}$ requires only an estimate of the number of termites in the colony.

Finally, Zimmerman *et al.*¹ estimate that there are 2.4×10^{17} termites in the world producing 151.6 Tg CH₄ yr⁻¹, or an average of 1.7 μg per day per termite; yet average emissions from the five colonies they studied¹ show emission rates of 0.2–0.6 μg per day per termite or 3–8 times less than required to achieve 150 Tg yr⁻¹. We agree that termites produce methane, however, at present we find no reason to believe that this source amounts to a worldwide production of 150 Tg yr⁻¹.

We thank Dr P. J. Fraser of CSIRO-Australia for helpful discussions. This work was supported in part by the Biospheric Research Corp. and the Andarz Co.

M. A. K. KHALIL
R. A. RASMUSSEN

*Department of Environmental Science,
Oregon Graduate Center,
Beaverton, Oregon 97006, USA*

1. Zimmerman, P. R., Greenberg, J. P., Wandiga, S. O. & Crutzen, *Science* **218**, 563–565 (1982).
2. Rasmussen, R. A. & Khalil, M. A. K. *Nature* (in the press).
3. Peakins, G. J., & G. Josens, in *Production Ecology of Ants and Termites*, (ed. Brian, M. W.) 111–164 (Cambridge University Press, 1978).
4. Haverly, M. I. *Pest Control* **44**, 12–17, 46–49 (1976).
5. Breznak, J. A. in *Symbiosis* 559–580 (Cambridge University Press, 1975).
6. Wood, T. G. in *Production Ecology of Ants and Termites*, (ed. Brian, M. W.) (Cambridge University Press, 1978).

Cooperation and competition in lions

IN a recent paper on lions¹ (*Panthera leo* L.), Packer and Pusey present valuable new data on competition between the members of male coalitions. I think, however, that without amplification their main conclusions may mislead the unwary.

Packer and Pusey state that "breeding coalitions of male lions include non-relatives much more commonly than was generally supposed". Their figure of 42%, that is, 5 out of 12 coalitions of known origin, is not in fact significantly different from the earlier figure of 10% (2 out of 21) quoted by Bygott *et al.*² Furthermore, as only the smaller coalitions included non-relatives, even in Packer and Pusey's sample 74% of males (29 out of 39) were with relatives; in the sample of Bygott *et al.*² the proportion was considerably higher. Because coalitions with relatives are larger (median size of 4 as opposed to 2), they retain tenure for about 2½ times as long² (thus fewer different large coalitions can be observed during the study period). Thus on average a lion potentially interacting with a coalition partner will be doing so with a relative in about 9 cases out of 10. There is no doubt that in either sample, most productive male companionship-years are spent with relatives.

Once companion males were engaged in competition with one another, Packer and Pusey detected no difference in the intensity of competition between related and unrelated coalition members; these are most interesting data. Should they be interpreted as indicating that kin selection is insignificant, or do they tell us something about how finely tuned natural selection is? If the optimal level of competition has evolved on the basis that the great majority of interactions are among relatives, and most of the successful breeding is by related males, do we necessarily expect that natural selection can also produce a different optimal level for the few cases of competition among non-relatives?

Previous ownership of an oestrous female is clearly of great importance; Packer and Pusey have contributed good quantitative data on that factor. As they point out, "game theoretical analysis predicts that when costs of fighting are high", (in relation to the benefits), "contests may be settled 'conventionally' through recognition of asymmetries such as 'owner versus rival'". The role of kin selection is that in competition between related males, it makes the costs higher and the benefits lower. It is certainly not the whole story. As I have pointed out³, the kin selection pressure is only one of the selection pressures influencing the costs and benefits of competing. Kin selection and game theory should not be regarded as in any sense alternative 'explanations' for low levels of aggression. I believe it would

be both a mistake and an over-reaction if kin selection were to be dismissed as insignificant in understanding male lion cooperation and competition. It is the most likely explanation for the complete absence of large coalitions containing non-relatives. Equally, we should not overemphasize its role. Almost certainly, it is important too in many other aspects of the social behaviour of lions as of other species.

My disagreement with some of Packer and Pusey's interpretations should emphasize the value of cooperation among unrelated lionologists in maintaining the long-term records which provide the basis for these excellent data and for the resulting discussion about them.

B. C. R. BERTRAM

*The Zoological Society
of London,
Regent's Park,
London NW1 4RY, UK*

1. Packer, C. & Pusey, A. E. *Nature* **296**, 740-742 (1982).
2. Bygott, J. D., Bertram, B. C. R. & Hanby, J. P. *Nature* **282**, 839-841 (1979).
3. Bertram, B. C. R. in *Growing Points in Ethology* (eds Bateson, P. P. G. & Hinde, R. A.) (Cambridge University Press, 1976).

PACKER AND PUSEY REPLY—We do not dismiss the significance of kin selection in the evolution of male cooperation in lions; kinship is obviously important in the choice of coalition partner¹. However, kinship is not necessarily a prerequisite for cooperation between male lions. Bygott *et al.*² showed that as coalition size increased, the individual reproductive success of each male in the coalition also increased. Therefore cooperative behaviour will be favoured by natural selection, even in the absence of kinship.

Bertram bases his explanation for the lack of difference in levels of competition between kin and non-kin on the assumption that only a very low proportion of interactions are between non-relatives. This raises two immediate questions. First, how should the proportion of males interacting with non-relatives be measured? Second, as the records are only now becoming long enough for us to estimate the proportion of males that have unrelated companions, just how rare are interactions between non-relatives?

We reported the percentage of coalitions containing non-relatives because this was the measure used by Bygott *et al.*². In estimating the likelihood of cooperating with non-relatives, a better measure might be the proportion of each male's partners that are non-relatives. Thus, each male in a coalition composed entirely of non-relatives cooperates with non-relatives 100% of the time, a male in a coalition with one related partner and one non-relative cooperates with a non-relative 50% of the time, and so on. By this measure, an average of 28% of each male's partners were non-relatives.

Our recent observations (April 1982) provide further evidence of the readiness with which males form coalitions with non-relatives. Of the 12 coalitions of known origins whose composition we reported, one pair of siblings has since been ousted from their pride and has joined with a third male from another pride. A trio of related males was reduced to one and the surviving male formed a coalition with an unrelated partner. Of the 39 males of known origins listed in our paper¹, 38% have so far had unrelated companions and 33% of each male's cooperative partnerships have been with unrelated companions.

Although most males probably start out with related companions, they are subjected to high mortality as subadults. By the time they are old enough to take over a pride, they may have lost one or more of their initial companions and replaced them with unrelated companions. Similarly, singleton males need time to find a partner after they have left their natal pride. The 21 coalitions of known origins reported by Bygott *et al.*² included coalitions of young males that had not yet gained a pride, whereas our sample of 12 coalitions was restricted only to those coalitions that had gained access to a pride. Five coalitions containing non-relatives out of 12 is considerably higher than 2 out of 21 ($P = 0.09$, two-tailed Fisher test), and, after including our most recent data, 7 of 12 is significantly higher than 2 of 21 ($P < 0.01$).

The existence of very large coalitions of relatives is an important indicator of recent levels of cub survival. The Serengeti lion population underwent a rapid increase in the 1970s³ and very large coalitions appeared at the same time. Previously, almost all male coalitions were pairs⁴. Our observations of the formation of coalitions of unrelated males suggest that males tolerate additional males only while they do not have access to a pride. The probability of a coalition gaining a pride increases rapidly as coalition size increases from one to two to three, and almost all coalitions of three gain a pride². Because males will not accept further partners while in control of a pride, coalitions of unrelated males are unlikely to 'grow' larger than three.

CRAIG PACKER
ANNE E. PUSEY

*Allee Laboratory of
Animal Behaviour,
University of Chicago,
940 East 57th Street,
Chicago, Illinois 60637, USA*

1. Packer, C. & Pusey, A. E. *Nature* **296**, 740-742 (1982).
2. Bygott, J. D., Bertram, B. C. R. & Hanby, J. P. *Nature* **282**, 839-841 (1979).
3. Hanby, J. P. & Bygott, J. D. in *Serengeti: Dynamics of an Ecosystem* (eds Sinclair, A. R. E. & Norton Griffiths, M.) (University of Chicago Press, 1979).
4. Schaller, G. B. *The Serengeti Lion* (University of Chicago Press, 1972).

In love with science

Donald MacKay

Realism and the Aim of Science. Vol. I of Postscript to the Logic of Scientific Discovery.

By Karl Popper, edited by W.W. Bartley III.

Hutchinson/Rowman & Littlefield: 1983. Pp.402. £20, \$35.

TO BECOME a legend in one's own time is evidently not an unmixed blessing — particularly if the legend is untrue to life. In this volume, the last to be published of the trilogy from his *Postscript to the Logic of Scientific Discovery*, Sir Karl Popper vigorously demolishes a number of "legends" (his own term) that have attached themselves to his controversial position over the past 50 years. He does so first in a tightly packed Introduction dated 1982, and then in a series of footnotes and addenda to the main text, most of which dates from 1956–1957.

The legend he is most anxious to refute is that "Popper is a naive falsificationist" (p.xxxiv) — i.e. someone who believes that an empirical theory can in practice be *conclusively* falsified by experiment. The refutation is simple — the citation of a string of passages from Popper's *Logic of Scientific Discovery* (1934) where he expressly denied that belief. He traces the trouble to a confusion of falsifiability-in-principle (which for Popper is an essential logical criterion of the scientific character of a statement) with *practical* falsifiability (which he has always recognized as problematic). "An entire literature", he argues, "rests on the failure to observe this distinction" (p.xxii). Once recognized, it makes nonsense of the further "legend" (p.xxxi) "that Thomas S. Kuhn . . . has shown that my views on science . . . can be refuted by . . . the history of science On the question of the significance of falsification for the history of science, Kuhn's and my views coincide almost completely". Extracts reproduced from Kuhn's treatment of Popper's views leave no doubt as to the justice of Popper's complaint. Many will feel it high time that pretentious claims for Kuhn as offering some novel "alternative to Popper" were cut down to size.

The first part of this volume (which logically forms an introduction to Vols II and III of the *Postscript*, reviewed in *Nature* 300, 663–664; 1982), is devoted to a sustained attack on what Popper calls "Inductivism" — the idea that valid scientific

information is gained by induction from observations. Against it, he spells out his well-known claim that only the refutation of hypotheses leads to the discovery of regularities:

My central idea in the field of animal knowledge (including human knowledge) is that it is based



on inherited knowledge All the information received from outside is eliminative, selective All knowledge remains fallible, conjectural. There is no justification, including, of course, no final justification of a refutation. Nevertheless, we learn . . . by refutations, i.e. by the elimination of errors, by feedback [p.xxxv].

Popper's general thesis is powerfully (and amusingly) illustrated by an analysis of Freud's notorious practice of turning all observations (even apparently contradictory ones) into "verifying instances" of his theory (pp.162–174). Against the objection that every falsification of a statement logically verifies its negation (so that

"falsification" and "verification" are symmetrical), Popper shows convincingly the logical asymmetry between falsification and verification of *theories*, pointing out that a theory may be falsified without any other (non-trivial) theory being *ipso facto* verified (p.186).

The Vienna Circle's adoption of "verifiability" as a criterion of "meaningfulness" in the 1920s was, of course, anathema to Popper, especially since "it (unintentionally) declared scientific theories to be meaningless" (p.175). But attempts by anti-metaphysicists to substitute "falsifiability" as a criterion are firmly rebutted:

The broad line of demarcation between empirical science on the one hand, and pseudo-science or metaphysics or logic or pure mathematics on the other has to be drawn right through the very heart of the region of sense — with meaningful theories on both sides of the dividing line . . . I reject . . . the dogma that metaphysics must be meaningless [pp.175–176, my italics].

Different concepts of probability are analysed at length in the concluding section. Popper's original preference (in *The Logic of Scientific Discovery*) for a "frequency" interpretation, which led to difficulties in assigning a probability to unique events, now gives way to the idea that relative frequencies of "repeatable" events (such as coin-tossings) are outward manifestations of a hidden objective physical disposition or "propensity" to produce them, which is the proper subject of judgements of probability, and which can be defined for singular as well as repetitive events.

The whole work shares with the other volumes of the *Postscript* an immense capacity to stimulate and clarify the mind of even the least sympathetic reader. My main criticism would be of the quite intemperate measure of "overclaim" that Popper seems to think necessary (perhaps for didactic reasons?) in spelling out his anti-inductivist principles. His assertion (p. xxxv, quoted above) that in learning "all the information received from outside is eliminative", for example, seems quite culpably to neglect the *instructive* role of "feedforward" from sense organs. While he admits that our guesses may be suggested to us by observations (p.13), and while he is quite right to draw attention to the element of active participation in seeing or perceiving (p.45) as against the "bucket theory of the mind", he has practically nothing to say about the ways in which the form of that participation is shaped by current input. The level of discussion here is typified by his astonishing claim (p.99) that "the fact that

a blind and deaf man may know more, and may make greater contributions to knowledge, than one with sharp eyes and ears" *refutes* "the traditional view that our knowledge grows by the accumulation . . . of perceptions or observations" (my italics).

Again, Popper's eagerness to depreciate inductive learning leads him to claim (p.98) that in science "all *new* knowledge, . . . all discoveries, are the result of trial and error" (rather than repetition). He even suggests (on p.50) that every author of an experimental paper should state in advance the specific hypothesis he intends to propose or discuss or test. This surely does less than justice to those phases of scientific investigation in which new territory has to be mapped before any non-trivial hypotheses can be formulated. Recording the first evoked brain potentials to a new stimulus, for example, may reveal nothing significant until one has made perhaps 100 repetitions; and what then emerges as an average out of the "background noise" may be a totally unexpected waveform which it would be merely impertinent and pretentious for the scientist to describe as a "test of his hypothesis".

The fact of the matter is that falsification of hypotheses is only one (albeit a highly desirable) way in which scientific information can be generated. True, we never start with a blank mind; but putting our eye to a new telescope, recording radio signals from distant stars or lifting a stone to see what lives underneath it are perfectly respectable ways of gaining scientific information even when we have only the vaguest of ideas as to what we may see. Until we have appropriate perceptual categories, we may indeed miss much that we may later learn to see in the same data; and unless we eventually find ourselves framing and testing hypotheses, our scientific enterprise will not have got very far. But in a new field especially, our scientific quest may have to begin by simply soaking ourselves for a time in raw data, letting our imagination play on them in the hope of gradually (or suddenly) spotting things to spur our curiosity and get the process of hypothesis-making and testing under way. Objective detachment in the scientific sense is entirely compatible with such imaginative commitment, since the object of the exercise is eventually to help all comers, whatever their values or attitudes, to reckon objectively with the reality we have encountered through it.

But let me not end on a negative note; for what distinguishes Popper from a great dull army of philosophers of science is that reading him is *good* for us. Even where he is (occasionally) outrageous he is at least sharp and clear enough for us to know better what we think in reaction; and at his best he is unsurpassed in the ability to recall to our minds the neglected emphases that make objective science worthy of our love. Yes, "love" is Popper's word for it:

I think that there is only one way to science — or

to philosophy, for that matter: to meet a problem, to see its beauty and fall in love with it; to get married to it, and to live with it happily, till death do ye part — unless you should meet another and even more fascinating problem, or unless, indeed, you should obtain a solution. But even if you do obtain a solution, you may then discover, to your delight, the existence of a whole family of enchanting though perhaps difficult problem children for whose welfare you may work, with a purpose, to the end of your days [p.8].

Donald MacKay, a physicist, is Emeritus Professor of Communication and Neuroscience at Keele University.

The view from Washington Square

Daniel D. Eley

Electronic Processes in Organic Crystals.

By Martin Pope and Charles E.

Swenberg.

Oxford University Press: 1982. Pp.821.

£75, \$145.

LYONS and Gutmann's *Organic Semiconductors* (Wiley, 1967) and Meier's book of the same title (Verlag Chemie, 1974) could reasonably be considered to present a comprehensive and balanced view of organic semiconductors. At that time, organic metals research was in its infancy, and organic superconductors still a gleam in the eye of W.A. Little. Any comparable treatment would these days need to account for at least 3,700 references, so Pope and Swenberg's disclaimer in their preface, that no attempt was made towards completeness nor objectivity in their choice of subject matter, must surely be understood.

How in fact is their book made up? The first four chapters, totalling 578 pages, deal with anthracene and its congeners, the last two with "materials of high dark conductivity" and "miscellaneous systems", a final ten-page postscript bringing us up to date on some nine topics such as solitons and superconductivity. At the outset one must record the general impression that the authors started with an enthusiastic exposition of their own main research topic of polycyclic aromatics, and only as an afterthought decided to include shorter reviews of some currently actively researched materials, inserting general theories where necessary; for example, the Schmidlin-Roberts statistical treatment of donors and acceptors in dark conduction first appears on p.662 in connection with phthalocyanines.

In the four "anthracene chapters" Pope and Swenberg systematically discuss from a theoretical viewpoint optical properties, charge carrier generation as single positive or negative carriers or as pairs in the bulk, and photoemission, and give an authoritative account of the literature relevant to

crystalline photoconductive molecular insulators. Complete beginners might find it helpful to first read Silinsh's introduction (*Organic Molecular Crystals*, published by Springer-Verlag in 1980) since there is a mass of detail and closely argued mechanisms to master. The book is largely free of typographical errors but one important one must be pointed out; in the energy level diagram on p.203, surely the valence and triplet state levels should be interchanged?

Turning to Chapter 5, we read about radical-ion salts and charge-transfer complexes, transitions in one-dimensional systems, (SN)_x and (TMTSF)₂PF₆ (the first organic superconductor at high pressure, updated to (TMTSF)₂ClO₄ with its ambient pressure superconductivity in the postscript). The following chapter includes phthalocyanines, polydiacetylenes, PVK and its TNF complex and various liquids. Perhaps rather strangely, the currently much-studied doped polyacetylene receives only three pages or so, with AsF₅ being first described as electron donor (p.700) and then as acceptor (p.701) — *bonus dormitat Homerus!*

Oxford University Press advertise this book as suitable for second-year graduate students and specialists, a recommendation I would endorse with the proviso that the former in their first year should follow up the earlier historical development of the subject in one or other of the two monographs mentioned earlier. In his foreword Sir Nevill Mott describes the book as "massive and comprehensive", whereas massive and selective seems a more correct description; an equally long volume would be needed to deal with a whole range of topics omitted from this work. The problem here is that it is often the minority topics which contain the germs of future successes. Mott goes on to opine that "it is indeed open to doubt, in view of the widely divergent explanations published of some phenomena, whether a universally agreed upon model will ever emerge".

Let us take heart from the fact that some 60 years separated the Van't Hoff-Le Bel tetrahedral carbon atom from the hybrid orbitals of Slater and Pauling, and a lot of valuable chemistry was carried out in the interval. It must be said that the need to set out so many physical theories has tended to obscure the simple physico-chemical guidelines which still help to direct organic synthesis to compounds with new electrical properties. Even if a unified physical theory must wait till the turn of the century, we may still hope for practical developments, useful materials to stand alongside PVK-TNF and justify our efforts. The present volume represents a valiant effort to bring order into an ever-expanding field, and help us all on our way.

Daniel D. Eley is Leverhulme Emeritus Fellow in the Department of Chemistry, University of Nottingham.

A matter of sex

John H. Crook

Primate Paradigms: Sex Roles and Social Bonds.

By Linda Marie Fedigan.

Eden Press, Montreal, and St Albans, Vermont: 1982. Pp.386. Pbk \$18.95.

IN evaluating *Primate Paradigms*, the orientation of the author and of her intended readers is of prime importance. Dr Fedigan writes about primate behaviour from an anthropological and sociological standpoint that differs greatly from the usual position of a biologist, whose concern with genes, mechanisms, sociobiological evolution and the origins of proto-cultures tends to preclude too questioning a reflection on the bias which his or her own background and sex may have on the ideas presented. It is upon this issue that Fedigan focuses in her quest to inform an audience of "non-specialists of all kinds, of students and of specialists in Primatology and Women's Studies".

In brief this is the feminist critique of theories of primate and human evolution, theories that are held to be in large part the "mythological" constructions of male observers. And the evidence is both copious and convincing. There can be no doubt that the date of a primatological publication, the *Zeitgeist* of the period so far as sex relations are concerned, the sex of the author, his or her social background and political orientation have all played some part in determining the acceptability of many popular viewpoints and of theory construction in primatology.

In a text of considerable competence and erudition, Fedigan examines many paradigms in primatological theory. Especially to the fore are hypotheses concerning dominance and alliance, kinship, sociosexual behaviour, the development of sex differences, sex roles in societies, sexual selection in evolution and theories of the origin of human social life. In many instances recent work does indeed demonstrate that the earlier literature (say in the 1960s and early 1970s) tended to overemphasize the roles, the activity, the dominance and the evolutionary influence of males, and to assign a general lack of importance to evolution of the female sex. In addition it can be argued that the species most studied tended to be exactly those which demonstrated such conditions most forcibly. As the sample of species studied has widened, so a greater awareness of such biases has emerged. Yet since this also coincided with the development of women's studies there can be little doubt that the unconscious prejudices of the most recent generation of field workers have changed from those of earlier ones. Whether the result has been a more objective primatology is a matter the author leaves in doubt

and which her personal bias would in any case prevent her from saying.

For all that *Primate Paradigms* is a notable book for it will alert primatologists and popularizers of science alike to the power of human projection especially in fields that bear on human action and social history. It is, however, important to note the limitations of such a text. Fedigan's approach is to ferret out all available contrary evidence to the paradigm in question and thereby to raise doubt regarding its integrity. She is of course not alone in recognizing the inadequacies of, for example, the early baboon-based models of human evolution and male workers have supplied more than a fair share of the critical arguments. Her persistent scepticism leads to chapters which repeatedly show their subject matter to contain so much contrary evidence that few conclusions can be reached.

Perhaps the generation of provocative hypotheses and their vigorous disproof prior to the erection of another is especially characteristic of the polemics of male academics. Few of them suppose their version to be the final truth, yet the promulgation of debate, the discovery of new facts and the weaving of new viewpoints is the very life-blood of a science. Fedigan almost never confronts a hypothesis directly nor does she seem to have any theory of her own. An erudition that merely casts doubts, avoids direct controversy and asserts a personal viewpoint is not one that will attract those who enjoy the more confrontational tactics common to much of biology. But Fedigan here is not writing as a biologist. She is

essentially a social critic of a branch of biological science.

Her argument is strongest where the evidence is most controversial, as in theories of social behaviour. Where she begins to overreach herself, and where her critical abilities are weakest, becomes apparent in her attitude to fundamental biological theories such as the origin of sexuality itself and contemporary research on sexual selection. Her cavalier treatment of the speculations of biologists is based on the assumption that male scientists are asserting directly or indirectly that women are inferior because eggs are bigger than sperm. This exceedingly trite view of the motives and views of leading and thoughtful researchers requires explanation in terms of ways in which contemporary women view their own collective identity, a social psychological issue clearly raised by Fedigan's controlled passion. She consistently underestimates the power of biological theory in the attempt to explain sexuality in the animal kingdom, and her dismissal of almost everything as "myth" seriously undermines her case. Biologists, especially cell biologists, are not likely to be impressed.

Nonetheless this is a courageous and hard-hitting book that should generate serious reflection. Its true position in the continuing debate about women in society will become clearer only when the influence of feminism on the judgement of women social scientists has been properly assessed.

John H. Crook is Reader in the Department of Psychology at the University of Bristol.

What is wrong with Muslim science?

Francis Ghiles

Science and Technology in the Middle East.

By Ziauddin Sardar.

Longman/Gale Research: 1982. Pp.322. £39, \$85.

AT ITS peak about one thousand years ago, the Muslim world made a remarkable contribution to science, notably mathematics and medicine. Baghdad in its heyday and southern Spain built universities to which thousands flocked: rulers surrounded themselves with scientists and artists. A spirit of freedom allowed Jews, Christians and Muslims to work side by side. Today all this is but a memory.

Expenditure on science and technology may have increased in recent years though that increase has been, perforce, limited to oil-rich countries: Sudan, for instance, is internally bankrupt and countries such as Iraq are busy fighting wars which cost billions of dollars — no doubt they have little time for science. Trade structures are dom-

inated by imported technology and most countries have economic and scientific systems geared to imitation rather than originality.

Even the recent wealth provided by oil exports makes relatively little difference. As the author of *Science and Technology in the Middle East* points out, "science policy and politics, much to the displeasure of many scientists, are closely linked in the Middle East". The region is dominated by dictatorships, benevolent or otherwise, the volatile nature of regimes such as those in Iran, Iraq, Syria and Libya further complicating any attempt to allow science to take root indigenously. Not surprisingly — and contrary to the author's belief — the brain drain to industrialized countries continues to debilitate intellectual life throughout the Middle East.

The book includes a 75-page "Overview" and a bibliography, and then gives a catalogue of regional organizations and their specific aims. The author also looks at 19 countries individually and provides a directory of major universities and research institutions. Compiling a reference book of this sort is an immensely complex task, one fraught with difficulties because few countries provide information which can be relied upon to be accurate, let

alone up to date. The data given here, however, seem to have been extracted from the dusty files of international organizations and are in places woefully inaccurate. Pakistan claims 34 pages, Turkey 23 and Egypt 15: dozens of institutes and research centres are listed yet one wonders how many actually function.

If one turns to Turkey, many more research projects on solid fuels and solar energy are under way than the author suggests. He is clearly impressed by Saudi Arabia and not very *au fait* with North Africa, otherwise he would not describe northern Algeria as having "thickly forested" mountains — nor would he suggest that some of the newer universities, such as Constantine, are more technically minded than older ones dominated by social sciences: Constantine is a real dump, where the Muslim fundamentalists have nearly wrecked the atmosphere necessary for real study in an institution of higher learning.

It is difficult to know how many inaccuracies the book contains as no observer can be familiar with every Muslim country — but the chapter on Egypt, for instance, is full of mistakes: it is simply not true to say that the Aswan High Dam has brought about a raised water table. That is rather the result of poor management, bad drainage of the canals and over-irrigation now that water is freely available; and if the author believes any figures for the reclamation of land in Egypt one can only wish him luck. He is also unreasonably harsh on the Gezina scheme in Sudan: its design is technically Victorian and for years it provided Sudan with the bulk of its foreign exchange.

In many ways the book reads like a series of hand-outs from universities and international organizations. It is out of date, riddled with errors and one gets the impression that the author has not set foot in some of the countries he describes — I cannot guess at the publishing strategy behind the book, nor can I see to whom it could be of any use. □

Francis Ghiles is a specialist writer for the Financial Times.

Pagels in Britain

Heinz R. Pagels' *The Cosmic Code: Quantum Physics as the Language of Nature* is now available in Britain. The UK publisher is Michael Joseph, price is £10.95. For review, see *Nature* 296, 503 (1982).

India's environment

The State of India's Environment 1982: A Citizens' Report (reviewed in *Nature* 302, 92; 1983) is available from Earthscan, 10 Percy Street, London W1P 0DR. Prices, which include postage and packaging, are £12.50 for Europe and \$22 for North America.

Corrigendum

In the review of three geology textbooks (*Nature* 302, 182; 1983) line 13 of the second column should have read "anchoring the geotherm".

Satellites, by Jove!

S.R. Taylor

Satellites of Jupiter.

Edited by David Morrison.

University of Arizona Press: 1982.

Pp.965. \$49.50.

"THE sense of novelty would probably not have been greater had we explored a different Solar System", remarked one observer after seeing the first fruits of the encounters of Voyager with the Jovian system. That sense of novelty is captured well in *Satellites of Jupiter*, a comprehensive account of what we have learned from the pictures sent back by both Voyager and Pioneer.

Contributors to the book concentrate mainly on the lunar-sized Galilean satellites, the 24 chapters providing a feast of information. Topics discussed range from the dark faint Jovian ring (showing that planetary rings are common, not rare) through orbital parameters, surfaces, cratering histories, geology, thermal evolution histories, and atmospheres, concluding with the origin and evolution of the system. Each account is written with the sophistication and expertise of post-Apollo planetary science, contrasting strongly with the literature of 20 years ago.

Much attention is given to the spectacular sulphur volcanism of Io, driven by tidal heating from Jupiter; the same phenomenon is probably responsible for resurfacing the thin icy crust of Europa. Ganymede, perhaps the most fascinating of the satellites, shows evidence of a few per cent expansion in radius following accretion. Melting of high density polymorphs of ice in the deep interior (about 50 per cent of Ganymede consists of water) and refreezing of the water at the surface as the familiar Ice I accounts for this expansion.

Cratering history is most clearly preserved on Callisto (resurfacing of Io and Europa has removed nearly all the craters). The population of objects which struck Callisto appears to differ from that in the inner Solar System, where the Moon, Mars and Mercury all record an early post-accretion bombardment by a similar suite. Still in dispute, however, is the question of whether all these heavily cratered surfaces are saturated, or reflect the total post-accretion impact history; most of the recent craters are thought to be due to cometary impact. Attempts to simulate cratering of an icy surface are described, but as recorded in Chapter 11 the problems of extrapolation are immense. (Incidentally, it is curious to note that one of the more successful target materials consisted of walnut shells, ground to sand size, lying on a clay slurry.)

Impressive though *Satellites of Jupiter* is, it is sad to note that the quality of

reproduction of many of the pictures — which provide so much of the primary information — is poor to the point of being useless. In addition the book inevitably represents a beginning rather than an ending. The new evidence is largely photographic, enabling the construction of geological maps, the establishment of cratering histories and so on, but we lack the compositional evidence to understand origins fully. This becomes clear in the final chapter on origin and evolution. The Galilean satellites show a decrease in density with distance from Jupiter — inferred as an increase in their content of water and other volatiles, probably due to an initial thermal gradient in a proto-Jupiter nebula. All the regular satellites are thought to have formed in this nebula on short-time scales (10^5 – 10^6 yr), while the irregular satellites were most likely captured.

We are thus closer than was Galileo to understanding the origin of the Jovian satellites. But one is reminded of the pre-Apollo debates over lunar origins — for the vital information we must await future missions. □

S.R. Taylor is a Professorial Fellow at the Research School of Earth Sciences, Australian National University, Canberra. His most recent book is Planetary Science: A Lunar Perspective (Lunar and Planetary Institute, 1982.)

The positive and the negative

Brian Coe

The Origins of Photography.

By Helmut Gernsheim.

Thames & Hudson: 1982.

Pp.280. £25, \$50.

PHOTOGRAPHY was introduced to an astonished world in 1839 by two individuals. One, L. J. M. Daguerre, was a French artist-showman who had perfected a process invented by his late predecessor Nicéphore Niépce. The Daguerreotype technique, the first to be announced, was out of general use within 20 years. It was replaced by other processes based on the discoveries of the Englishman W. H. F. Talbot, whose Photogenic Drawing process was announced soon after that of Daguerre.

Talbot was a very different character from Daguerre; he was a talented mathematician, a botanist of some distinction, a classical scholar and noted Assyriologist. But above all he was an experimental scientist whose painstaking and systematic researches led to the discovery in 1835 of a photographic process in which a negative was produced, from which unlimited numbers of copies could be made. Modern photography has

evolved directly from Talbot's work.

Helmut Gernsheim's book sets out to document these early years of photography, up to the early 1850s when new processes made both Daguerre's and Talbot's pioneering methods obsolete. The text is substantially the same as that of the first part of *The History of Photography* by Alison and Helmut Gernsheim, published by Thames & Hudson in 1969. This original work, although by no means a complete account of the evolution of photography and often displaying the prejudices of the authors, was nonetheless a valuable assembly of information on photographic history, published at a time when there was little literature available (and, it must be said, relatively little interest in the subject). In the past 13 years all that has changed — a considerable number of books have appeared on the subject, many of them based on original research. This new edition takes little account of this fresh body of knowledge. For example, Gernsheim does not appear to have examined the notebooks and correspondence of Talbot held at the Science Museum in London and at Lacock Abbey, and his hostile assessment of Talbot is based largely on secondary and often biased sources. The publication of H. J. P. Arnold's *William Henry Fox Talbot*, a definitive biography of Talbot by Hutchinson-Benham in 1977, has set the record straight, but no account is taken of this in Gernsheim's book.

In a work the size of the original, it was understandable that some errors should have occurred, but few of these have been corrected in the new edition. For example, Gernsheim repeats the story that Glasshouse Street, off Regent Street in London, was so named because of the number of photographic studios to be found there. A good story — but Glasshouse Street appears in rate-books as early as 1678, and is shown on a map of London published in 1681. And the statement that the cost of a Daguerreotype outfit, £16 in 1839, "does not seem high compared with the prices of modern precision cameras" was silly when

it was made in the first edition, and is even sillier today after 13 years of inflation and the dramatic fall in the real cost of photographic equipment. These examples are typical of a number of mainly minor inaccuracies which appear mostly in the author's comments on the usually well-documented facts.

At first sight the book, with its large page size (10 by 11 inches) and 191 illustrations, is impressive, but with many of the illustrations an error of judgement has been made. The Daguerreotype process produced an image in a whitish silver-mercury amalgam on a highly reflective silver surface. The image was seen as a positive only when the plate was held so that the mirror-like silver surface reflected a dark ground. In this book, the illustrations of Daguerreotypes are printed in black ink on a silver ink ground. While it is true that this gives the reproduction a metallic effect, the impression it produces is precisely the opposite to that seen in a real Daguerreotype. What is worse, the reflective ink often makes the picture difficult to see if the lighting is at the wrong angle. A few full-colour reproductions of tinted Daguerreotypes printed without the silver backing show the nature of the image far more successfully. The absence of an index, except for a list of names, is a serious omission in a book of such pretensions, although space was found for a biography of the author which would have been more at home on the dust jacket.

Gernsheim's *History of Photography* has been a standard work of reference for many years; it is a pity that its shortcomings were not corrected in this new version. Nonetheless, in its new form the book is still a substantial work of reference which, if used with discretion, will provide a valuable source of information about the early days of a medium which plays such an important part in modern life. □

Brian Coe is the Curator of the Kodak Museum at Harrow, Middlesex. Among other books he is the author of Cameras, from Daguerreotype to Instant Picture (Nordbok, 1977).

Behaviour, body and brain

Euan M. Macphail

Primate Brain Evolution: Methods and Concepts.

Edited by Este Armstrong and Dean Falk.
Plenum: 1982. Pp. 330. \$39.50, £27.65.

AS THE primates evolved they grew bigger, as did their brains; those larger brains (to judge from living primates) showed more subdivisions than did the earlier, smaller brains. Why did primate brain size increase, and why are there apparently novel brain areas in more advanced primates? The 19 chapters in this volume, which represent contributions to two symposia, are in general brief progress reports on a variety of approaches to these questions.

One obvious reason for the increase in brain size is that a larger body requires more brain space for processing somatic information. But can increasing somatic demand be quantified, so that the contribution of other factors (such as increasing intelligence) to increases in brain size may be assessed? Jerison argues that it can and that somatic demands on brain volume vary in accordance with body surface area. Other contributors challenge this conclusion: in the chapters by Radinsky, Martin, and Holloway and Post we find that the mean rate of increase of brain size with body weight (used by Jerison to deduce somatic demand) for a group of vertebrates depends heavily on the nature of the sample used.

If it is hard to agree upon a useful measure of relative brain size, it is equally difficult to identify a given brain structure as "new" and to show that it is associated with a novel behavioural capacity. Campbell introduces the problem of identifying homologies in the nervous system, and several contributors assess potentially novel features found in

IMAGE
UNAVAILABLE FOR
COPYRIGHT
REASONS

advanced primates; other chapters, having a similar logic, concern changes in organization or in relative size of structures found throughout primates. Although some speculations on behavioural relevance are introduced, no convincing case is made for any of the proposed links of brain to behaviour.

The book as a whole falls uncomfortably between two stools: it provides very few new data for those familiar with the area, but does not form a comprehensive survey of major topics. Most of all, it lacks a detailed consideration of behavioural data.

Uncertainty principle of theology

John Habgood

Science and the Renewal of Belief.

By Russell Stannard.

SCM Press, 56-58 Bloomsbury Street,
London: 1982.
Pp.207 Pbk £2.95.

IN matters of religious belief it is not uncommon to find theologians in some ways more sceptical than scientists. No doubt the converse is also true. Those professionally engaged in a subject are likely to be more aware of its internal strains and inconsistencies, and the unanswered questions, than those who look at it as a whole from the vantage point of some other discipline.

Professor Stannard has a foot in both camps. A high-energy physicist, and one of the discoverers of charm in fundamental particles, he is also a Reader in the Church of England, who has found the developments in his own science a reinforcement to his faith. He uses his scientific knowledge effectively in tackling some of the major problems put to him by would-be believers, and for those who are stuck on such questions as evolution and the Bible, the reductionist view of human nature, miracles, or the alleged persecution of Galileo, his book has much to offer. I particularly like a discussion of the problem of evil, which draws on Dirac's concept of anti-matter to make the point that things have to be defined negatively as well as positively. His chatty style, frankness when he feels he has to offend some of the more simple Christian believers, and general non-nonsense approach to the subject, make this an attractive package which I am reluctant to criticize.

And yet. . . In the end I am left with the impression that his would-be believers did not press him hard enough. It is too easy, for example, to claim that a serious experiment with prayer can answer the question of God's existence, in much the same way that a scientific experiment can verify the existence of charm. There is an analogy, certainly, but it sidesteps the

Primate brains are of interest primarily because we are primates, with remarkable intellectual capacities: in order to be sure that any of the brain features we examine are of general interest, we need to know that they are in fact related to intellectual activity, and that can be established only by considering behaviour alongside neurology. □

Euan M. Macphail is Senior Lecturer in Psychology at the University of York. He is author of Brain and Intelligence in Vertebrates (Clarendon, 1982).

major philosophical and theological problems in a way which should be impossible for somebody writing from within these disciplines. What seems to be lacking is the sense of mystery, and of theology's attempt to struggle with what is ultimately inexpressible.

There is a revealing passage right at the end of the book where Professor Stannard writes, "An advantage of being a scientist is that this inclination to sit on the fence is not one to which we are particularly prone. . . Among the ranks of scientists are to be found atheists and Christians, but very few waverers in the middle". Precisely. Straightforward questions must have clear answers. The essence of science is so to arrange experience that clear answers can be given. But what if life as it is actually lived does not, and cannot, have the same kind of clarity? And what if God has to be found as much in the fog of uncertainty as in the light of faith? □

John Habgood has been Bishop of Durham since 1973. He lectured in physiology and pharmacology at the University of Cambridge before ordination.

Extra dimensions to globins

Warner E. Love

Haemoglobin & Myoglobin. Atlas of Molecular Structures in Biology, Volume 2.

By G. Fermi and M.F. Perutz.
Oxford University Press: 1982. Pp.104.
£22.50, \$49.50.

PROTEIN molecules are complex, intricately assembled three-dimensional objects. Different aspects of a particular molecule's structure will be of interest to different investigators, and in the end it may be that only three-dimensional models can suffice to fulfil particular needs. In the case of haemoglobin and myoglobin, structures have been determined at near-atomic resolution of molecules from different species, with different amino acid sequences and in different ligand states.

In this book G. Fermi and M.F. Perutz have brought together an enormous amount of data from widely scattered sources, spread over some 40 years of study, and have provided an excellent summary of our current knowledge of haemoglobin and myoglobin. The book is thus a good reference basis for further study of the literature, and should also be helpful to teachers faced with the problem of presenting an introduction to our understanding of how proteins work, i.e. how their structures subserve their physiological functions. In this regard, myoglobin and haemoglobin are good models; their structures are very well known and there is a large body of physiological and physical chemical data available for interpretation in terms of their structures.

In all cases thus far examined, those molecules that reversibly bind oxygen at an iron porphyrin binding-site incorporate that site into a polypeptide chain of approximately 17,000 MW, folded as is sperm whale myoglobin. Vertebrate haemoglobins are tetramers composed of two alpha and two beta myoglobin-folded chains. In the first part of the book, Fermi summarizes the relevant physiological phenomena, for example oxygen binding and cooperativity, and the influence of allosteric effectors such as CO₂, pH and organic phosphates. The amino acid sequences and intra-subunit hydrogen bonds of human and horse alpha and beta chains, and of sperm whale myoglobin, are given, the sequences are aligned and the helices designated. The contacts between subunits for human deoxy and horse met haemoglobins are illustrated in order to display the alterations in those contacts that occur upon the quaternary structure transition that accompanies ligand binding.

The descriptive material in Fermi's section sets the stage for 35 red-green stereodiagrams. These illustrate salient details of the three-dimensional structures, namely, skeleton diagrams of myoglobin and haemoglobin, their haem pockets, subunit interfaces and horse met haemoglobin helices.

In the second part of the book Perutz discusses fetal and adult mutant human haemoglobins from the point of view of how their deviations from the standard physiological behaviour of normal adult human haemoglobin can be rationalized in terms of the anomalies seen in their crystal structures. To date crystal structures of 20 different mutant haemoglobins have been determined, and the success with which their altered physiologies are explained in structural terms is one of the triumphs of modern X-ray crystallography.

I have already found this book useful, both in teaching and in my research. I have no doubt that others will find it equally valuable. □

Warner E. Love is Professor in the Thomas C. Jenkins Department of Biophysics at the Johns Hopkins University.